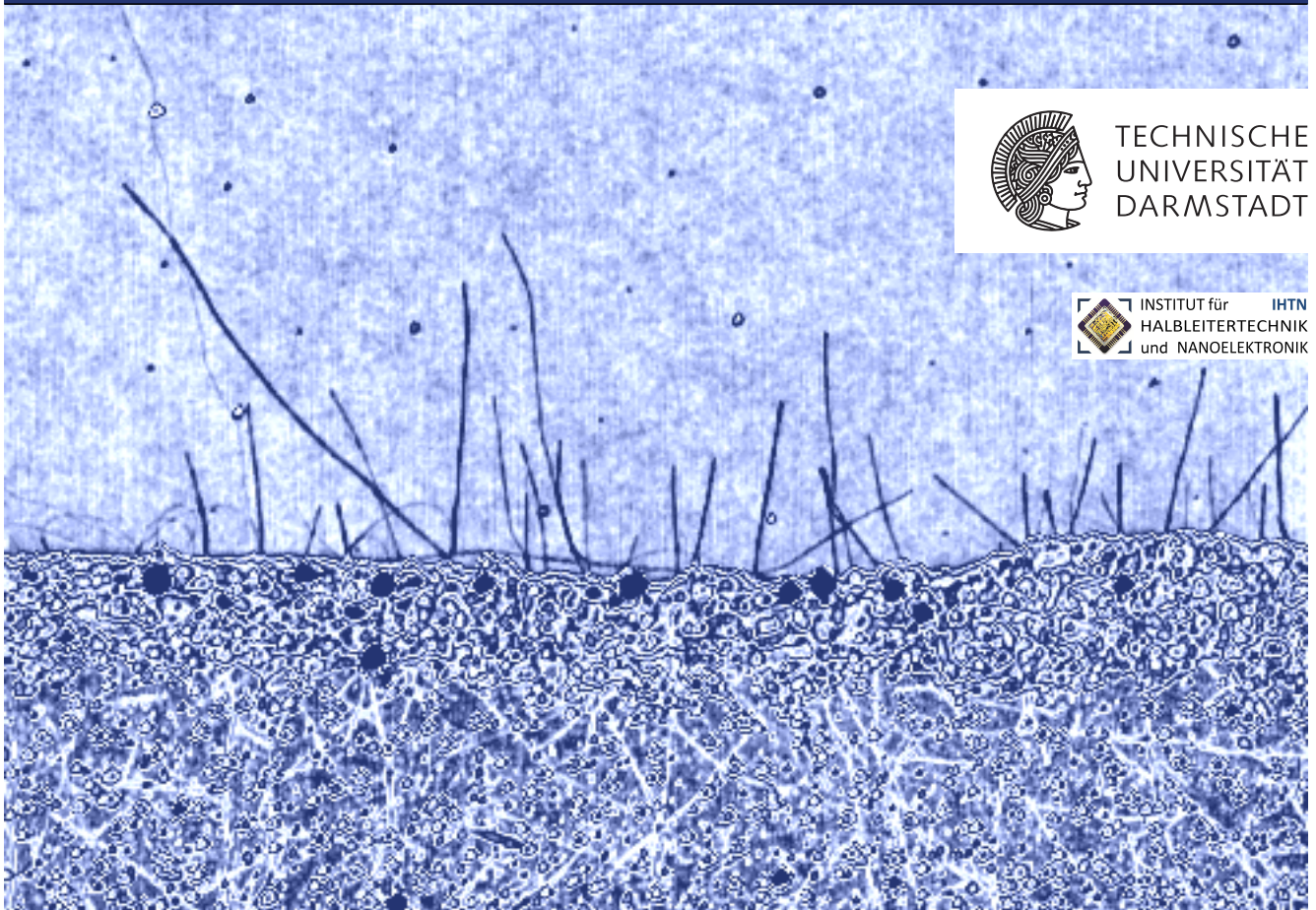


Untersuchung einer Herstellungstechnologie für Feldeffekt-Transistoren auf Basis von Kohlenstoffnanoröhren

**Investigation of a manufacturing technology
for carbon nanotube based field-effect transistors**

Zur Erlangung des akademischen Grades Doktor-Ingenieur (Dr.-Ing.)
genehmigte Dissertation von Dipl.-Ing. Martin Keyn aus Grimma
Tag der Einreichung: 26.01.2018, Tag der Prüfung: 20.06.2018
Darmstadt — D 17

1. Gutachten: Prof. Dr. rer. nat. Udo Schwalke
2. Gutachten: Prof. Dr.-Ing. Thomas Mikolajick



TECHNISCHE
UNIVERSITÄT
DARMSTADT



INSTITUT für IHTN
HALBLEITERTECHNIK
und NANOELEKTRONIK

Untersuchung einer Herstellungstechnologie für
Feldeffekt-Transistoren auf Basis von Kohlenstoffnanoröhren
Investigation of a manufacturing technology
for carbon nanotube based field-effect transistors

Genehmigte Dissertation von Dipl.-Ing. Martin Keyn aus Grimma

1. Gutachten: Prof. Dr. rer. nat. Udo Schwalke
2. Gutachten: Prof. Dr.-Ing. Thomas Mikolajick

Tag der Einreichung: 26.01.2018

Tag der Prüfung: 20.06.2018

Darmstadt — D 17

Bitte zitieren Sie dieses Dokument als:

URN: urn:nbn:de:tuda-tuprints-76606

URL: <http://tuprints.ulb.tu-darmstadt.de/7660>

Dieses Dokument wird bereitgestellt von tuprints,
E-Publishing-Service der TU Darmstadt
<http://tuprints.ulb.tu-darmstadt.de>
tuprints@ulb.tu-darmstadt.de



Die Veröffentlichung steht unter folgender Creative Commons Lizenz:
Namensnennung – Weitergabe unter gleichen Bedingungen 4.0 International
<https://creativecommons.org/licenses/by-sa/4.0/>



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Untersuchung einer Herstellungstechnologie für Feldeffekt-Transistoren auf Basis von Kohlenstoffnanoröhren

Vom Fachbereich Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades eines
Doktor-Ingenieurs (Dr.-Ing.)

genehmigte Dissertation

von

Dipl.-Ing. Martin Keyn

geboren am 15. Juli 1984 in Grimma

Referent:	Prof. Dr. rer. nat. Udo Schwalke
Korreferent:	Prof. Dr.-Ing. Thomas Mikolajick
Tag der Einreichung:	26.01.2018
Tag der mündlichen Prüfung:	20.06.2018

D 17
Darmstadt 2017



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Erklärung laut §9 PromO

Ich versichere hiermit, dass ich die vorliegende Dissertation allein und nur unter Verwendung der angegebenen Literatur verfasst habe. Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

(Martin Keyn)

Darmstadt,
den 26.01.2018

Kurzfassung

Seit der Erfindung der integrierten Schaltung in den 1960er Jahren wird die Skalierung elektronischer Bauelemente dem Postulat von MOORE folgend immer weiter vorangetrieben. Die Skalierung des wichtigsten Bauelements der Mikro- und Nanoelektronik, des Metall-Oxid-Halbleiter-Feldeffekt-Transistors (MOSFET), wird dabei in absehbarer Zeit an Grenzen stoßen, da unter anderem durch zunehmende Kurzkanal-Effekte und Leckströme der Zugewinn der Skalierung mehr und mehr reduziert wird. Das ultimative Limit für die Skalierung stellt letztendlich der unterhalb von etwa 5 nm Abstand zwischen *Source* und *Drain* einsetzende direkte Tunnelstrom dar. Technologische Hindernisse, Probleme der Zuverlässigkeit und erschwerte Prozessbeherrschbarkeit sowie insbesondere die stets zwingend erforderliche Wirtschaftlichkeit der Technologieerweiterungen verlangsamen zusätzlich die Fortführung der Silizium-basierten Skalierung. Die Entwicklung neuer Ansätze und alternativer Technologien ist daher gefragt. Eine solche Technologie stellt der Kohlenstoffnanoröhren-Feldeffekt-Transistor (CNTFET) dar, der in Form einer Hybrid-Technologie auf Basis von Silizium und Kohlenstoff realisiert werden kann. Der CNTFET ersetzt dabei das Silizium im Bereich des Kanals, also den aktiven, über den Feldeffekt gesteuerte Teil des Transistors, durch halbleitende einwandige Kohlenstoffnanoröhren (SWNTs).

Eine Herstellungstechnologie für CNTFETs ist am INSTITUT FÜR HALBLEITERTECHNIK UND NANOELEKTRONIK (IHTN) verfügbar. Diese nutzt das *in situ*-Wachstum von CNTs auf Basis der katalytisch-chemischen Gasphasenabscheidung (CCVD), wobei Methan als Kohlenstoffquelle verwendet wird. Der metallische Aluminium/Nickel-Katalysator wird zuvor durch eine Temperung in Stickstoff-Atmosphäre in nanoskalige Nickel-Cluster und ein dielektrisches Aluminiumoxid transformiert („*sacrificial catalyst*“). *Source/Drain*-Kontakte aus Palladium werden mittels der *Lift off*-Technik direkt auf die gewachsenen CNTs strukturiert. Die resultierenden Bauelemente zeigen unter Nutzung der Waferrückseite als globale *Backgate*-Elektrode unipolares p-MOS-Verhalten.

In dieser Arbeit werden auf Basis des vorhandenen Aluminium/Nickel-Katalysators Variationen von Prozessparametern der Temperung und des CCVD-Schritts mit dem Ziel durchgeführt, möglichst viele und möglichst dünne SWNTs zu wachsen. Der Einfluss auf die Cluster-Bildung und das CNT-Wachstum wird durch topologische Untersuchungen mittels Rasterkraftmikroskopie (AFM) bewertet. Dabei können hohe Heizraten der Temperung, eine Cluster-Bildungs-Temperatur von 920 °C, die Verwendung von Argon anstelle von Stickstoff

als Inertgas und eine Sauerstoff-Plasmabehandlung des Katalysators als zielführend ermittelt werden.

Ein Testchip-Layout und ein modifizierter Katalysator werden im Verlauf der Arbeit eingeführt. Dieser Katalysator nutzt die Atomlagenabscheidung (ALD) zur direkten Herstellung einer Aluminiumoxid-Schicht, die als Untergrund für die Nickel-Schicht dient. Das Testchip-Layout ist in der Größenordnung von Mikrometern gehalten, wobei die vorgesehenen Transistoren nominale Kanalweiten im Bereich von Millimetern aufweisen. Durch die großen Kanalweiten soll eine elektrische Parallelisierung von CNTs ermöglicht werden, um eine Skalierung des *On*-Stroms der Transistoren zu erreichen.

Auf Basis des neuen Testchip-Layouts, des modifizierten Katalysators und des bekannten CCVD-Prozesses wird auf einem Substrat mit 100 nm Trocken-Oxid eine statistische Untersuchung an über 1.500 Bauelementen durchgeführt. Die untersuchten Transistoren weisen nominale Kanalweiten von 20 μm und 470 μm sowie Längen von 4 μm auf. Diese Untersuchung ergibt eine Korrelation zwischen der nominalen Kanalweite und der Ausbeute an funktionalen Bauelementen bzw. deren Schaltverhalten. Für die große Kanalweite werden hohe Ausbeuten, aber nur geringes Schaltverhalten der Transistoren beobachtet. Für die kleine Kanalweite dagegen wird eine schlechtere Ausbeute erzielt, jedoch weisen die funktionalen Bauelemente ein hohes Schaltverhalten auf. *On*-Ströme von bis zu 3 μA bei $V_{DS} = -500\text{ mV}$ und ein Schaltverhalten mit einem *On/Off*-Verhältnis von bis zu sieben Größenordnungen können für beide nominalen Kanalweiten beobachtet werden. Hierbei beträgt die geringste ermittelte Unterschwellensteigung 120 mV/dec. Die mit dem neuen Testchip-Layout angestrebte Skalierung des *On*-Stroms kann derzeit somit bestenfalls eingeschränkt nachgewiesen werden.

Abstract

Since the invention of the integrated circuit in the 1960's, the scaling of electronic devices has been pushed further and further by obeying MOORE's paradigm. This will reach its limits in the near future for the most important device in micro and nano electronics, i. e. the metal-oxide-insulator field-effect transistor (MOSFET), due to short channel effects and leakage currents degrading scaling gains. The ultimate limit for scaling will be reached at a distance of approximately 5 nm between source and drain where direct tunneling will occur. Technological pitfalls, issues of reliability, demanding process control as well as the always essential profitability of technological advances will hinder further silicon-based scaling. Therefore new concepts and alternative technologies are urgently needed. Such a technology could come in shape of the carbon nanotube field-effect transistor (CNTFET), which can be fabricated based on a hybrid technology of silicon and carbon. The CNTFET substitutes the silicon channel, i. e. the active, via field-effect controlled part of the transistor, with a semi-conducting single-walled nanotube (SWNT).

A technology to fabricate CNTFETs is available at the INSTITUTE FOR SEMICONDUCTOR TECHNOLOGY AND NANOELECTRONICS (ISTN). It utilizes *in situ* growth of CNTs by means of catalytic chemical vapor deposition (CCVD) and uses methane as carbon source. The metallic aluminum/nickel catalyst is prior annealed in nitrogen to form nano-scaled clusters and a dielectric aluminum oxide ("sacrificial catalyst"). Palladium source/drain contacts are then directly patterned onto the grown CNTs using the lift off technique. The resulting devices demonstrate unipolar p-MOS behavior utilizing the wafer backside as a global backgate electrode.

In this work, on the basis of the existing aluminum/nickel catalyst, variations of process parameters of the annealing and the CCVD step are carried out with the aim to grow as many and as thin as possible SWNTs. The effect on the cluster formation and CNT growth is evaluated by topological investigations by means of atomic force microscopy (AFM). Here, increased heat rates for annealing, a cluster formation temperature of 920 °C, the utilization of argon instead of nitrogen as inert gas and a oxygen plasma treatment of the catalyst are found to be productive.

Furthermore, a new test-chip layout and modified catalyst are developed. The catalyst utilizes atomic layer deposition (ALD) to directly grow a layer of aluminum oxide which serves as a underlayer for the nickel. The layout's critical dimensions are in the micrometer range and transistors with nominal channel widths in the millimeter range are designed. Those great channel widths shall enabled electrical parallelization of CNTs to make scaling of the transistor *On-currents* possible.

Using the new test-chip layout, the modified catalyst and the known CNT growth process on a substrate with 100 nm of silicon oxide a statistical study is carried out on over 1,500 devices. Those transistors have nominal channel widths of 20 μm and 470 μm , respectively, and lengths of 4 μm . The study reveals a correlation of nominal channel width and the yield of functional devices as well as their switching behavior. For the big channel width a high yield of functional devices is obtained, but those devices only show poor switching behavior. On the contrary, for the small channel width only a low yield is obtained, but those devices show good switching behavior. *On*-currents of up to 3 μA at $V_{DS} = -500\text{ mV}$ and switching with an on/off ration of up to seven orders of magnitude of current can be observed for both nominal channel widths. Here, the minimal subthreshold slope has a value of 120 mV/dec. The scaling of the *On*-current which is intended by the new test-chip layout is only very limitedly reached at this point.

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen zu Kohlenstoffnanoröhren	3
2.1	Struktur	3
2.2	Elektronische Eigenschaften	5
2.3	Verfahren zur Herstellung	12
3	Grundlagen zu Feldeffekt-Transistoren	17
3.1	Metall–Halbleiter-Kontakt	17
3.2	MOS-Kondensator	20
3.3	MOS-Feldeffekt-Transistor	22
4	Analytische Methoden	25
4.1	Ellipsometrie	25
4.2	Rasterkraftmikroskopie	28
4.3	Elektrische Charakterisierung	32
5	Verwendete Prozesstechnologien	33
5.1	Reinigungsverfahren	33
5.2	Thermische Oxidation	35
5.3	Photolithographie	36
5.4	Strukturübertragung durch nasschemisches Ätzen	39
5.5	Strukturübertragung mittels <i>Lift off</i> -Technik	40
5.6	Elektronenstrahlverdampfung	43
5.7	Chemische Gasphasenabscheidung	44
5.8	Atomlagenabscheidung	46
5.9	Reaktives <i>Sputtern</i>	50
5.10	Kurzzeit-Prozessierung	51
5.11	Sauerstoff-Plasmabehandlung	52
5.12	Ozongenerator	53

6	Prozess zur Herstellung von CNTFETs	55
6.1	Testchip-Layout	56
6.2	Prozessablauf	60
7	Topologische Untersuchung von Cluster-Bildung und CNT-Wachstum	67
7.1	Beschreibung des CNT-Wachstumsmechanismus	67
7.2	Beschreibung des Basisprozesses	69
7.3	Einfluss der Sauerstoff-Plasmabehandlung des Katalysators	70
7.4	Einfluss der CNT-Wachstumsdauer	72
7.5	Einfluss der Cluster-Bildungs-Temperatur	74
7.6	Einfluss der Heizrate der Cluster-Bildung	76
7.7	Zusammenfassung der Ergebnisse und optimierte Prozessführung	83
8	CNTFETs mit Al/Ni-Katalysator	85
8.1	Prozess zur Herstellung der FETs	85
8.2	Elektrische Untersuchung von ausgewählten FETs	86
8.3	Vergleich der untersuchten FETs	89
9	Untersuchung des dielektrischen Versagens der CNTFETs mit Al/Ni-Katalysator	93
9.1	Lokalisierung der dielektrischen Durchbrüche	93
9.2	Mechanismus der Durchbrüche	95
9.3	Ursache des Versagens	96
10	Evaluierung anderer Katalysatoren	99
10.1	Schichtstapel der Katalysatoren	100
10.2	CNT-Wachstum auf den untersuchten Katalysatoren	102
11	CNTFETs mit Al₂O₃/Ni-Katalysator	105
11.1	Prozess zur Herstellung der FETs	105
11.2	Statistische Untersuchung der Ausbeute und des Schaltverhaltens	108
11.3	Untersuchungen von ausgewählten FETs	115
11.4	Elektrische Zerstörung metallischer CNTs	117
11.5	Bewertung der Ergebnisse	120
12	Zusammenfassung	123

Anhang	125
Herleitung des CNT-Durchmessers	127
Vollständiges Testchip-Layout	129
Basisparameter der Herstellung	133
Haftungsproblem der Palladium-Kontakte	135
Wafer-Map zu Kapitel 11.2	137
Schaltplan des Impuls-Generators	139
 Verzeichnisse	 141
Abbildungen	143
Tabellen	147
Abkürzungen	149
Symbole	151
Farbcodierung	155
Literatur	157
Eigene Veröffentlichungen	165
Betreute studentische Arbeiten	169
 Danksagung	 171
 Lebenslauf	 173

1 Einleitung

Seit der Herstellung und Patentierung des ersten integrierten Schaltkreises im Jahr 1958 durch KILBY [1,2] hat sich die Mikroelektronik ausgehend von einem einfachen handgefertigten Prototypen zu einem hochkomplexen Technologiefeld entwickelt. Die sogenannte Skalierung ist dafür die technologische Triebfeder, die die Mikroelektronik schon längst hat zur Nanoelektronik werden lassen. „Skalierung“ bedeutet dabei die kontinuierliche Verkleinerung der geometrischen Abmessungen der Bauelemente — sowohl in vertikaler als auch in lateraler Ausdehnung. Dem Postulat von MOORE folgend, der im Jahr 1965 eine Verdoppelung der Bauelementanzahl pro Chip alle 12 Monaten voraussagte [3], werden für das wichtigste Bauelement, den Metall-Oxid-Halbleiter-Feldeffekt-Transistor (MOSFET), mittlerweile physikalische Grenzen erreicht. Um den MOSFET dennoch weiterhin nutzen und verbessern zu können, werden neue Bauelementkonzepte wie FinFETs oder Silizium-Nanodrähte [4,5] entwickelt. Auch die Entwicklung einer Kohlenstoff-basierten Elektronik, also der Einsatz bisher in der Halbleitertechnologie nicht genutzter Materialien, wird verfolgt. Hierbei sind der Kohlenstoffnanoröhren-FET (CNTFET) und der Graphen-FET (GFET) zwei der wichtigsten Vertreter [6,7].

Die erste, in Abb. 1a gezeigte Aufnahme von CNTs der sowjetischen Wissenschaftler RADUSHKEVICH UND LUKYANOVICH datiert bereits auf das Jahr 1952 [8]. Aufgrund des Kalten Kriegs und der vorherrschenden Sprachbarriere zum Westen fand ihre Arbeit allerdings kaum Beachtung. Deswegen wird die eigentliche Entdeckung der CNT rund vierzig Jahre später dem Japaner IJIMA für seine 1991 veröffentlichte Arbeit zugeschrieben [9], die die wohl bekannteste Aufnahme von CNTs enthält, die hier in Abb. 1b gezeigt ist. Dass Elektronik auf Basis von CNTs tatsächlich möglich ist, wurde 1998 von TANS, VERSCHUEREN UND DEKKER demonstriert, die einen Transistor realisierten, der eine CNT als Kanal nutzt [10].

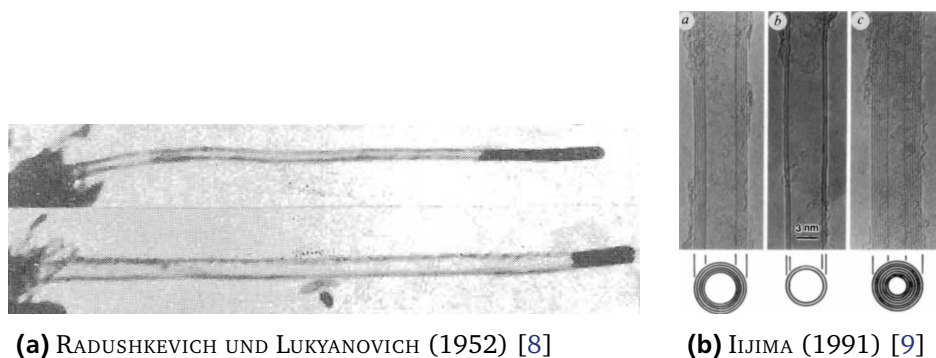


Abb. 1: Erste Aufnahmen von Kohlenstoffnanoröhren (CNTs).

Ziel dieser Arbeit ist es, den am INSTITUT FÜR HALBLEITERTECHNIK UND NANOELEKTRONIK (IHTN) und in der Vorgängerarbeit [6] entwickelten technologischen Ansatz weiterzuverfolgen und zu vertiefen. CNTs werden hierbei *in situ* unter Verwendung der katalytisch-chemische Gasphasenabscheidung (CCVD) gewachsen, wobei der Katalysator in Form von nanoskaligen Clustern vorliegt, die zuvor durch eine Temperung gebildet werden. Verschiedene Parameter dieser beiden Prozesse sollen untersucht werden, um besonders den Einfluss der Cluster-Bildung auf die später gewachsenen CNTs zu untersuchen, um möglichst lange und möglichst dünne einwandige Kohlenstoffnanoröhren (SWNTs) zu erhalten. Auch soll hier ein neues Testchip-Layout zum Einsatz kommen, das große nominale Kanalweiten aufweist. Durch die weiten Kanäle soll eine elektrische Parallelisierung von CNTs erreicht werden, sodass Transistoren mit höheren *On*-Strömen hergestellt werden können.

Gliederung der Arbeit

Diese Arbeit gliedert sich im Wesentlichen in zwei Teile: in den Grundlagen-Teil, der direkt an diese Einleitung anschließt, und den Ergebnis-Teil, auf den seinerseits eine Zusammenfassung dieser Arbeit und ab Seite 125 der Anhang und die verschiedenen Verzeichnisse folgen.

In den Grundlagen werden in Kapitel 2 zunächst die festkörperphysikalischen Eigenschaften von CNTs und die möglichen Herstellungsverfahren referiert. Kapitel 3 behandelt in Kürze die Grundlagen zu Feldeffekt-Transistoren, die für die Interpretation der hier gemachten elektrischen Messung nötig sind. Die genutzten analytischen Verfahren werden in Kapitel 4 dargestellt. Besonderes Augenmerk soll dabei auf der Rasterkraftmikroskopie (AFM) und speziell ihren Einschränkungen liegen, um Messartefakte zu erkennen und die Ergebnisse hinsichtlich der Cluster-Bildung und des CNT-Wachstums zutreffend bewerten zu können. Abgeschlossen wird der Grundlagen-Teil mit einer Beschreibung der hier verwendeten Prozesstechnologien in Kapitel 5.

Der Ergebnis-Teil beginnt mit Kapitel 6, in dem der entwickelte Prozess zur Herstellung von Feldeffekt-Transistoren auf Basis von CNTs behandelt und das entwickelte Testchip-Layout diskutiert wird. In Kapitel 7 werden anhand topologischer Untersuchungen mittels AFM die Auswirkungen verschiedener Parameter der Prozessführung untersucht, die während der Temperung und des CCVD-Schritts Einfluss auf die Cluster-Bildung und auf das CNT-Wachstum nehmen. Kapitel 8 und Kapitel 11 sind der elektrischen Charakterisierung von CNTFETs unter Verwendung des Aluminium/Nickel- bzw. des Aluminiumoxid/Nickel-Katalysators gewidmet. Dazwischen sind in Kapitel 9 und Kapitel 10 eine Untersuchung des dielektrischen Versagens der CNTFETs mit Aluminium/Nickel-Katalysator bzw. eine Evaluierung anderer Katalysatoren notwendig. Dieser zweite Teil der Arbeit wird mit einer Zusammenfassung der Ergebnisse in Kapitel 12 abgeschlossen.

2 Grundlagen zu Kohlenstoffnanoröhren

2.1 Struktur

Das Element Kohlenstoff bildet sechs verschiedenen Konfigurationen aus [11], die auch nach ihrer „Dimensionalität“ bezeichnet werden können. Mit Ausnahme von Diamant und Carbin (und leichten Einschränkungen bei den Fullerenen) können sie von 2-dimensionalen, unendlich ausgedehnten Schichten von wabenförmig angeordnetem Kohlenstoff abgeleitet werden.

- **3-D — Diamant und Graphit:**

Diamant und Graphit sind die beiden üblichen „volumenbehafteten“ Formen, wobei Graphit aus einzelnen Kohlenstoff-Schichten besteht und Diamant in der nach ihm benannten kubischen Struktur mit vierfach koordinierten Atomen aufgebaut ist.

- **2-D — Graphen:**

Graphen besteht nur noch aus einer oder wenigen Schichten. Es besitzt daher näherungsweise keine Höhe und somit eine Ausdehnung in nur zwei Raumrichtungen. Es wurde 2004 durch NOVOSELOV UND GEIM entdeckt [12].

- **1-D — Kohlenstoffnanoröhren und Carbin:**

CNTs können als „aufgerolltes Graphen“ beschrieben werden. Die Röhren dehnen sich nur in eine Raumrichtung aus. Entdeckt wurden sie 1991 durch IJIMA [9] bzw. — unbeachtet — bereits 1952 durch RADUSHKEVICH UND LUKYANOVICH [8]. Carbin dagegen besteht aus Ketten von Kohlenstoff-Atomen und wurde 1967 von KASATOCHKIN ET AL. zuerst beschrieben [13]. Ein natürliches Vorkommen von Carbin, dessen mechanische Eigenschaften denen von CNTs überlegen sein sollen [14], wurde 1982 von SMITH UND BUSECK allerdings als unwahrscheinlich und die bisherigen Nachweise als Fehlinterpretationen eingestuft [15]. Carbin wäre demnach die einzige Konfiguration des Kohlenstoffs, die nicht natürlich auftritt, sondern künstlich synthetisiert werden muss [16].

- **0-D — Fulleren:**

Das Buckminster-Fulleren C_{60} und höhere Fullerene sind Kugeln oder Käfige aus Atomen des Kohlenstoffs, die durch den geringen Durchmesser als quasi-punktförmig angesehen werden und somit keine räumliche Ausdehnung haben. Sie wurden 1985 durch KROTO ET AL. entdeckt [17].

Die Struktur von CNTs kann von Graphen abgeleitet werden, wobei eine einwandige Kohlenstoffnanoröhre (SWNT) dabei durch Aufrollen einer Graphen-Ebene entlang des Chiralitätsvektors \vec{C} entsteht. Werden mehrere Graphen-Ebenen gleichzeitig aufgerollt oder werden SWNTs mit unterschiedlichen Durchmessern ineinander geschachtelt, so entsteht eine mehrwandige Kohlenstoffnanoröhre (MWNT).

Der Chiralitätsvektor \vec{C} wird über die Komponenten n und m , die Gittervektoren \vec{a}_1 und \vec{a}_2 und einen Koordinatenursprung in Bezug auf das Graphen-Gitter gemäß Gl. 1 definiert. Damit berechnet sich der Durchmesser d_{SWNT} einer SWNT gemäß Gl. 2 über den Typ (n, m) und den Abstand $a_{\text{C-C}}$ zwischen zwei benachbarten Kohlenstoff-Atomen im Graphen-Gitter, wobei der Umfang U_{SWNT} der Röhre über die Länge $|\vec{C}|$ des Chiralitätsvektors gegeben ist. (Für die Herleitung von Gl. 2 siehe Anhang, Seite 127.) Theoretische Berechnung von SAWADA UND HAMADA zeigen, dass SWNTs mindestens einen Durchmesser von 0,4 nm haben müssen [18]. Typische Durchmesser liegen im Bereich von 0,7 bis 3 nm, wobei SWNTs mit einem Durchmesser von 1,4 nm energetisch am stabilsten sind.

$$\vec{C} = n \cdot \vec{a}_1 + m \cdot \vec{a}_2 \quad (1)$$

$$d_{\text{SWNT}} = \frac{U_{\text{SWNT}}}{\pi} = \frac{|\vec{C}|}{\pi} = \frac{\sqrt{3}a_{\text{C-C}} \cdot \sqrt{n^2 + nm + m^2}}{\pi} \quad (2)$$

Wie Abb. 2 zeigt, werden die Gittervektoren so gewählt, dass sie einen Winkel von 60° einschließen und der Koordinatenursprung auf einem der Kohlenstoff-Atome liegt [19]. Diese heute übliche Definition weicht dabei leicht von der ursprünglichen Definition von HAMADA ET AL. ab, die den Mittelpunkt einer der Sechsecke als Koordinatenursprung und Gittervektoren mit einem 120° -Winkel nutzt [20]. Es zeichnen sich zwei besondere Konfigurationen aus: die *zigzag*- und die *armchair*-CNT. Für beide leitet sich die Bezeichnung von der in Abb. 2a und Abb. 2b skizzierten Form ab, die sich entlang des Chiralitätsvektors an den Kanten der Sechsecke abzeichnet. Allgemein sind damit alle CNTs vom *armchair*-Typ, wenn für sie $n = 0$ oder $m = 0$ gilt, bzw. vom *zigzag*-Typ, wenn $n = m$ gilt. Alle übrigen Konfigurationen wie die in Abb. 2c skizzierte (2, 4)-CNT, die nicht einem dieser beiden Grenzfälle entsprechen, werden als „chiral“ bezeichnet. Als Sammelbegriff für *armchair*- und *zigzag*-CNTs wird manchmal auch die Bezeichnung „achiral“ verwendet.

Analog zum Chiralitätsvektor kann über den Typ (n, m) einer CNT ein Translationsvektor \vec{T} gemäß Gl. 3 definiert werden. \vec{T} steht im Graphen-Gitter senkrecht auf \vec{C} und zeigt damit in Richtung der Röhrenachse. Wichtiger als die Richtung von \vec{T} ist seine Länge gemäß Gl. 4, da so die Länge der Elementarzelle einer CNT bestimmt werden kann. Der Ausdruck vereinfacht sich für achirale CNTs, da er für diese unabhängig vom Chiralitätstyp (n, m) wird und schlicht über die Gitterkonstante des Graphen-Gitters a bzw. über die Bindungslänge $a_{\text{C-C}}$ ausgedrückt werden kann.

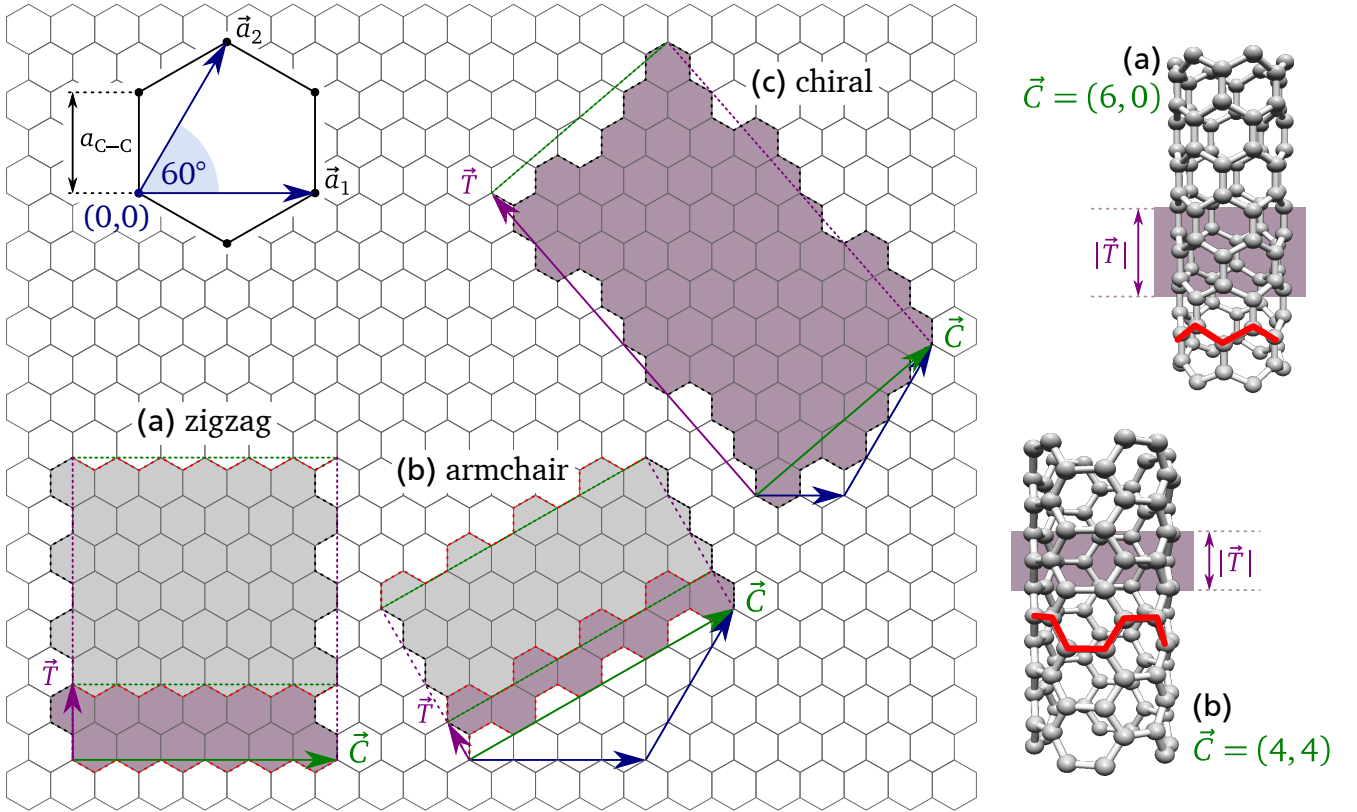


Abb. 2: Modellhafte Erklärung der Struktur von CNTs: Durch Aufrollen einer Graphen-Ebene entlang des Chiralitätsvektors \vec{C} (Gl. 1) entstehen für $n = 0 \vee m = 0$ zigzag- und für $n = m$ armchair-CNTs. Die Länge einer Elementarzelle der CNT ist über die Länge des Translationsvektors $|\vec{T}|$ gegeben (Gl. 4).

$$\vec{T} = -\frac{2m+n}{NR} \cdot \vec{a}_1 + \frac{2n+m}{NR} \cdot \vec{a}_2 \quad (3)$$

$$\text{wobei } N = \text{gcd}(n, m) \quad \text{und} \quad R = \begin{cases} 3 & \text{wenn } (n-m)/3N \in \mathbb{Z}_0 \\ 1 & \text{wenn } (n-m)/3N \notin \mathbb{Z}_0 \end{cases}$$

$$\Rightarrow |\vec{T}| = \frac{\sqrt{3(n^2 + nm + m^2)}}{NR} \cdot a = \begin{cases} \sqrt{3}a = 3a_{C-C} & \text{für zigzag} \\ a = \sqrt{3}a_{C-C} & \text{für armchair} \end{cases} \quad (4)$$

2.2 Elektronische Eigenschaften

Ebenso wie die Struktur können auch die elektronischen Eigenschaften der CNTs von Graphen abgeleitet werden. Daher sollen hier zunächst knapp die Eigenschaften von Graphen hergeleitet und beschrieben werden und schließlich die Schlussfolgerungen für das elektronische Verhalten von CNTs präsentiert werden. Eine umfassendere Darstellung findet sich unter anderen in [6, 21, 22] und begleitend in den Lehrbüchern der Festkörperphysik (z. B. [23]).

Atomorbitale von Kohlenstoff

Werden Atome als frei betrachtet, so können ihre Elektronen nur diskrete Werte annehmen, die als (Energie-)Zustände bezeichnet werden. Welche Zustände ein Elektron annehmen kann und mit welcher Wahrscheinlichkeit es dies tut, wird durch die Quantenmechanik beschrieben, die den Elektronen sowohl Eigenschaften von Wellen als auch von Teilchen zuordnet. Das Atomorbital Ψ ist dabei die räumliche Wellenfunktion eines Elektrons in einem bestimmten Zustand. Die Aufenthaltswahrscheinlichkeit des Elektrons an einem bestimmten Ort \vec{r} ist als Dichtefunktion über das Betragsquadrat $|\Psi(\vec{r})|^2$ gegeben. Oft wird mit Hilfe einer Isofläche von $|\Psi(\vec{r})|^2$ ein kugel- oder keulenförmiges Volumen beschrieben, das veranschaulicht, wo sich das Elektron mit hoher Wahrscheinlichkeit befindet.

Das Element Kohlenstoff besitzt sechs Elektronen, die die Orbitale 1s, 2s und 2p besetzen können. Zwei der Elektronen sind Kernelektronen, die das Orbital 1s besetzen. Zwei der vier Valenzelektronen befinden sich im 2s-Orbital, während die zwei restlichen Elektronen jeweils eines der drei 2p-Orbitale besetzen. Ein einzelnes Kohlenstoffatom im Grundzustand hat also die in Abb. 3a skizzierte Elektronenkonfiguration $1s^2 2s^2 2p^2$. Der angeregte Zustand, den Kohlenstoff im Kristallverbund einnimmt, ist in Abb. 3b gezeigt, wobei eines der Atome aus dem 2s-Orbital in das dritte 2p-Orbital angehoben wird; es sind nun vier einfach besetzte Orbitale vorhanden, die Bindungen zu anderen Atomen eingehen können. Ausgehend von diesem angeregten Zustand kann es zur Hybridisierung, d. h. Vermischung, der Orbitale kommen. Abhängig davon wie viele der drei 2p-Orbitale sich mit dem einen 2s-Orbital vermischen, wird von einer sp -, sp^2 - oder sp^3 -Hybridisierung gesprochen. Im Fall von Graphen kommt es zur sp^2 -Hybridisierung der Orbitale 2s, $2p_x$ und $2p_y$, die in Abb. 3c skizziert ist. Die drei Elektronen in den sp^2 -Orbitalen bilden dabei zu drei benachbarten Atomen sogenannte σ -Bindungen aus. Da jede der drei Bindungen für sich den maximalen Raum beansprucht, kommt es zu dem 120° -Winkel zwischen den Bindungen, der die im vorherigen Kapitel beschriebene hexagonale, wabenförmige Struktur des Graphen bedingt. Wie auch in Abb. 3d skizziert, spannen die σ -Bindungen in x - und y -Richtung eine Ebene auf, zu der in z -Richtung die $2p_z$ -Orbitale jeweils nebeneinander senkrecht stehen und daher als π -Bindungen bezeichnet werden. Da sie weniger stark sind, bilden sie ein „ π -Netzwerk“ mit delokalisierten Elektronen, das für die Leitfähigkeit von Graphen verantwortlich ist, während die σ -Bindungen den atomaren Zusammenhalt durch kovalente Bindungen herstellen.

Bänderstruktur von Graphen

Während einzelne Atome nur diskrete Zustände annehmen können, interagieren Atome in einem Festkörper miteinander. Diese Interaktion führt zu einer Multiplikation der Zustände, sodass abschnittsweise quasi-kontinuierliche Verläufe möglicher Zustände entstehen, die als (Energie-)Bänder bezeichnet werden. Am absoluten Nullpunkt der Temperatur, wenn ein Fest-

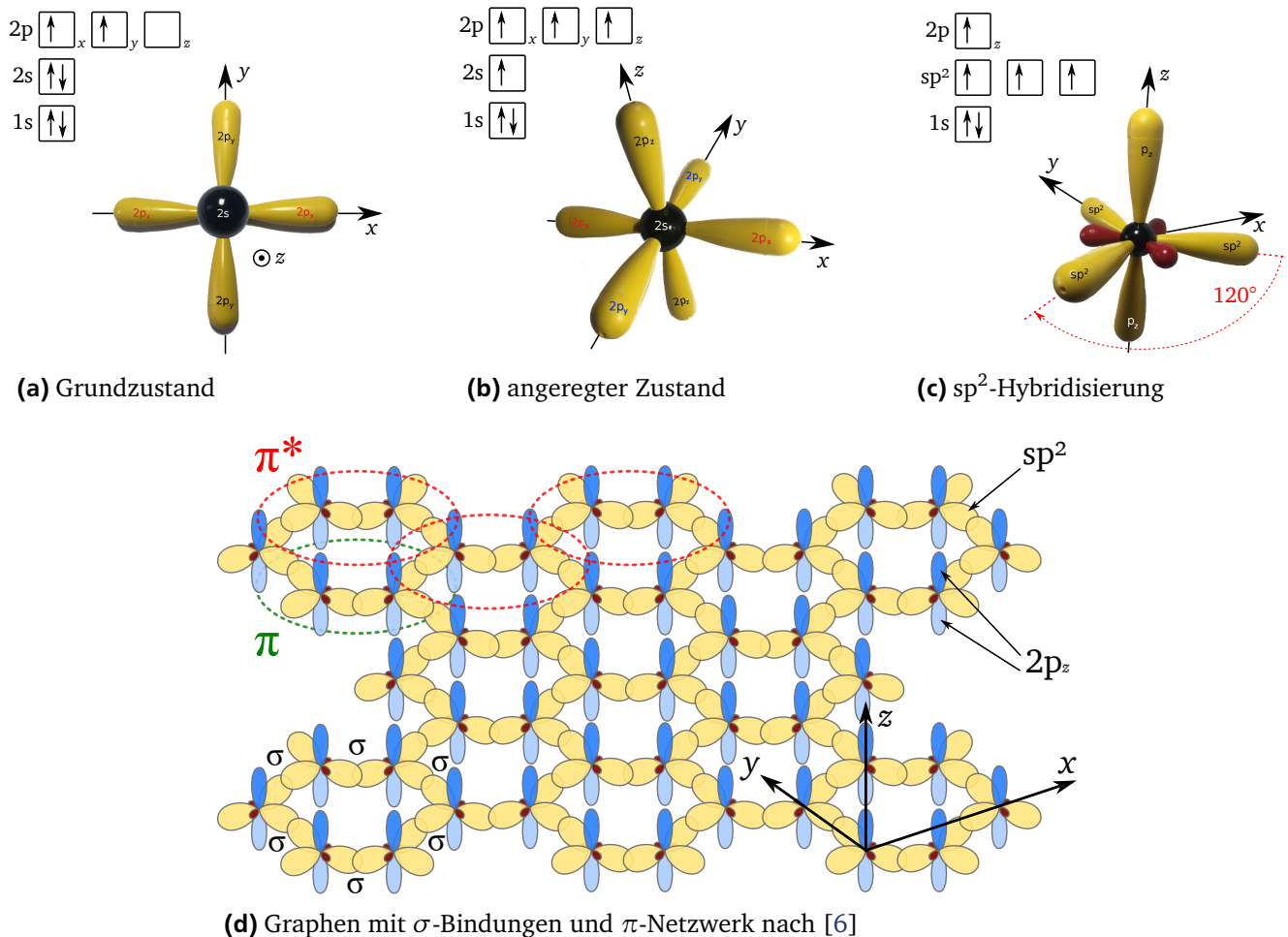


Abb. 3: Atomorbitale von Kohlenstoff: Ausgehend vom Grundzustand $1s^2 2s^2 2p^2$ wird ein Elektron aus dem 2s-Orbital in das dritte, noch freie 2p-Orbital in den angeregten Zustand angehoben. Davon ausgehend kommt es im Fall von Graphen zur sp^2 -Hybridisierung, d. h. der Vermischung der Orbitale 2s, $2p_x$ und $2p_y$. In Graphen überlappen sich die sp^2 -Orbitale und bilden kovalente σ -Bindungen, während die $2p_z$ -Orbitale zu einem π -Netzwerk mit delokalisierten Elektronen führen. (Abbildungen der Keulen-Modelle in (a) bis (c) aus [24].)

stoff seinen niedrigsten energetischen Zustand hat, sind die Bänder von unten nach oben mit Elektronen gefüllt. Dabei wird das Band mit der höchsten Energie, das zumindest noch teilweise mit Elektronen gefüllt ist, als Valenzband bezeichnet, das Band mit der niedrigsten Energie, das vollständig leer ist, als Leitungsband. Die Energiedifferenz zwischen der Unterkante des Leitungsbands und der Oberkante des Valenzbands ist die Bandlücke E_g . Liegt der Wert von E_g im Bereich von etwa 1 eV handelt es sich um einen Halbleiter; ist der Wert größer als 6 eV, ist es per Definition ein Isolator. Überlappen sich die beiden Bänder liegt ein Metall vor.

Da Elektronen Fermionen sind, folgen sie der Fermi-Dirac-Verteilung nach Gl. 5, die die Wahrscheinlichkeit f beschreibt, nach der ein energetischer Zustand E von einem Elektron besetzt ist. Dabei ist T die absolute Temperatur, k_B die Boltzmann-Konstante und E_f die Fermi-Energie. Der Fermi-Energie kommt besondere Bedeutung zu, da sie zum einen statistisch denjenigen ener-

getischen Zustand beschreibt, der bei $T \geq 0 \text{ K}$ mit genau 50 % Wahrscheinlichkeit von einem Elektron besetzt ist, und zum anderen in Halbleitern den Bezugspunkt für die Austrittsarbeit darstellt.

$$f_{\text{Fermi-Dirac}}(E) = \frac{1}{\exp\left(\frac{E-E_f}{k_B T}\right) + 1} \quad (5)$$

Zur Beschreibung der Bänderstruktur werden Elektronen in der Quantenmechanik als Wellen betrachtet, die durch ihren Wellenvektor \vec{k} im reziproken Raum beschrieben werden. Mit Hilfe der zeitunabhängigen Schrödinger-Gleichung ist es dann möglich, die erlaubten energetischen Zustände E in Abhängigkeit von der „Position“ \vec{k} im reziproken Raum zu berechnen. Eine direkte Darstellung ist üblicherweise nicht möglich, da Festkörper in der Regel dreidimensional sind und deswegen auch \vec{k} drei Komponenten aufweist (k_x, k_y, k_z), die gezeichnet werden müssen — was eine vierdimensionale Zeichnung zur Folge hätte. Stattdessen wird die Energie entlang bestimmter Linien innerhalb der Elementarzelle im reziproken Raum, der ersten Brillouin-Zone, skizziert, was einer Reduzierung auf eine Raumrichtung entspricht und eine zweidimensionale Darstellung ermöglicht.

Für Graphen ist der Zusammenhang zwischen realem und reziprokem Raum in Abb. 4a skizziert. Nach ASHCROFT UND MERMIN stehen die Einheitsvektoren des realen Raumes $\vec{a}_{1,2}$ dabei gemäß Gl. 6 paarweise senkrecht auf den Einheitsvektoren des reziproken Raums $\vec{b}_{1,2}$, wobei δ_{ij} das Kronecker-Delta ist [23]. Für die Konstruktion der Wigner-Seitz-Zelle wird ein Kohlenstoff-Atom der Wabenstruktur samt eines nächsten Nachbarn als Einheit betrachtet, sodass sich die grau markierte Parallelogrammform ergibt. Im reziproken Raum hat die orange markierte erste Brillouin-Zone drei unterschiedliche ausgezeichnete Punkte: die M-Punkte (Seitenmittelpunkte), den Γ -Punkt (Mittelpunkt) und die K-Punkte (Ecken); aufgrund der auch im reziproken Raum vorhandenen hexagonalen Struktur treten die M- und K-Punkte dabei jeweils sechsfach auf. Entlang einer Linie M– Γ –K–M (blau in Abb. 4) wird die Brillouin-Zone aufgefaltet, um die in Abb. 4b gezeigte zweidimensionale Darstellung der Bänderstruktur von Graphen zu erhalten. Durch die zweidimensionale Natur von Graphen kann allerdings die Raumkomponente k_z entfallen, was eine vollständige Darstellung der Bänderstruktur möglich macht, die in Abb. 4c abgebildet ist.

$$\vec{b}_i \cdot \vec{a}_j = 2\pi \delta_{ij} \quad \text{mit} \quad \delta_{ij} = \begin{cases} 0 & \text{wenn } i = j \\ 1 & \text{wenn } i \neq j \end{cases} \quad (6)$$

Wie Abb. 4b zeigt, kreuzen sich an den K-Punkten das π -Valenzband und das π^* -Leitungsband von Graphen. Wird der ganze \vec{k} -Raum betrachtet, bedeutet das, dass es sich bei Graphen weder um einen Halbleiter noch um ein Metall handelt, da weder eine Bandlücke noch eine Überlappung zwischen bzw. von Valenz- und Leitungsband vorliegt; Graphen wird daher

als „Halbmetall“ bezeichnet. Da zwischen den beiden σ -Bändern mindestens ein Abstand von 11 eV liegt, tragen sie kaum zum elektrischen Verhalten von Graphen bei und werden daher in viele Betrachtungen nicht mit einbezogen [21]. Für die Beschreibung der π -Bänder wird das *tight binding*-Modell verwendet, das bereits im Jahr 1947 von WALLACE für die ersten Berechnungen der Bänderstruktur von Graphen genutzt wurde [25]. Für eine quantitative Berechnung der Bänderstruktur von Graphen muss für alle Elektronen die Schrödinger-Gleichung nach Gl. 7 gelöst werden, wobei \hat{H} der Hamiltonoperator und E und Ψ die Eigenwerte und -vektoren des Wellenvektors \vec{k} sind.

$$\hat{H} \Psi(\vec{k}) = E(\vec{k}) \Psi(\vec{k}) \quad (7)$$

Auf die Lösung von Gl. 7 wird hier verzichtet und auf die bereits zitierte Literatur verwiesen. Festzuhalten ist, dass Graphen ein Halbmetall ist, dessen elektrische Leitfähigkeit auf den nicht hybridisierten $2p_z$ -Orbitalen und dem resultierenden π -Netzwerk von delokalisierten Elektronen beruht, und dass die „leitfähigen Punkte“ im reziproken Raum die Ecken (K-Punkte) der ersten Brillouin-Zone sind.

Elektronisches Verhalten von CNTs

Zur Herleitung der elektronischen Eigenschaften von SWNTs ist die Methode des *zone folding* gut geeignet, da mit ihr unabhängig vom konkreten Durchmesser und Chiralitätstyp der CNT die Bandstruktur bestimmt werden kann. Um einen Zusammenhang zwischen den elektronischen Eigenschaften und dem Chiralitätstyp (n, m) herzustellen, sollen der Wellenvektor entlang des

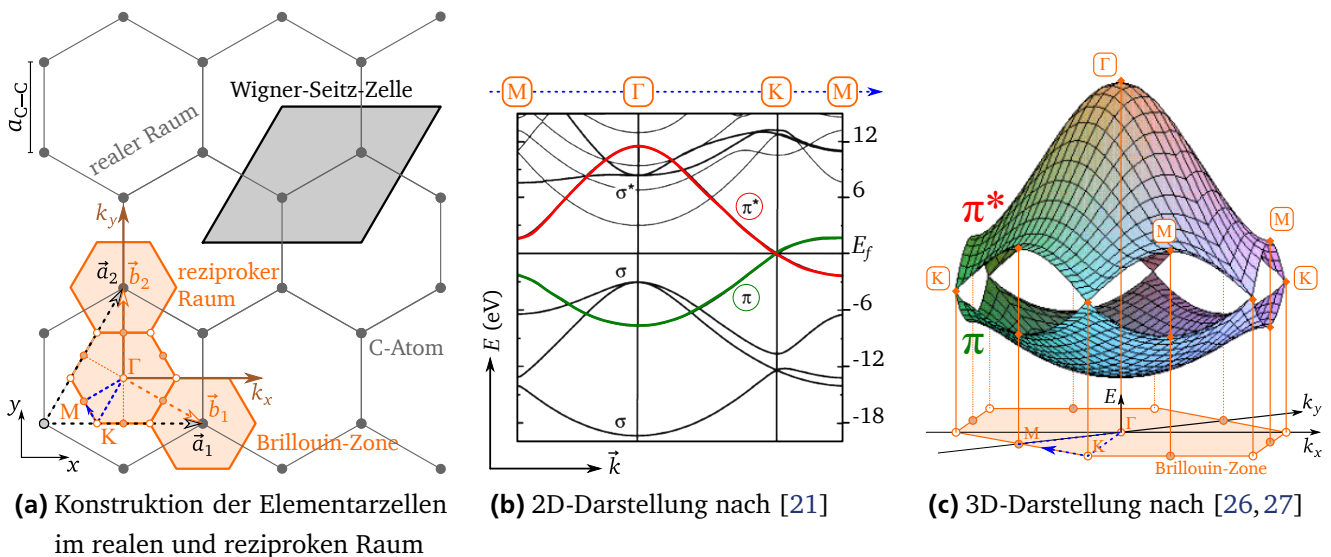


Abb. 4: Konstruktion und Darstellung der ersten Brillouin-Zone und Bänderstruktur von Graphen. Die σ -Bänder haben mindestens einen Abstand von 11 eV und tragen daher kaum zur Ladungstransport bei. Das π -Valenzband und π^* -Leitungsband kreuzen sich in den K-Punkten. Da weder eine durchgehende Bandlücke noch ein Überlappung der Bänder vorliegt, wird Graphen als Halbmetall bezeichnet.

Röhrenumfangs \vec{k}_\perp und entlang der Röhrenachse \vec{k}_\parallel senkrecht bzw. parallel zum Chiralitätsvektor \vec{C} stehen:

$$\begin{aligned} \vec{k}_\perp \cdot \vec{C} &= 2\pi & \text{bzw.} & & \vec{k}_\perp \cdot \vec{T} &= 0 \\ \vec{k}_\parallel \cdot \vec{C} &= 0 & \text{bzw.} & & \vec{k}_\parallel \cdot \vec{T} &= 2\pi \end{aligned} \quad (8)$$

Aufgrund der extremen Aspektverhältnisse kann die Röhre entlang ihrer Achse als unendlich ausgedehnt angesehen werden; das bedeutet, dass der Wellenvektor \vec{k}_\parallel kontinuierliche Werte annehmen kann. Die Länge k_\parallel des reziproken Gittervektors in Richtung der Achse ist dabei mit der Länge $|\vec{T}|$ der Elementarzelle der CNT verknüpft:

$$k_\parallel = \frac{2\pi}{|\vec{T}|} \implies k_\parallel \in \left(-\frac{\pi}{|\vec{T}|}, \frac{\pi}{|\vec{T}|} \right] \quad (9)$$

Der Wellenvektor \vec{k}_\perp dagegen ist quantisiert, weil aufgrund des geschlossenen Umlaufs entlang des Röhrenumfangs die Bedingung für eine stehende Welle erfüllt werden muss. Dies ist nur dann der Fall, wenn sämtliche Wellenfunktionen einen Phasenunterschied von ganzzahligen Vielfachen p von 2π aufweisen, wobei p mit der Anzahl q der Graphen-Sechsecke in der Einheitszelle der CNT verknüpft ist:

$$k_{\perp,p} = \frac{2\pi}{|\vec{C}|} \cdot p = \frac{2}{d_{\text{SWNT}}} \cdot p \quad \text{wobei} \quad p = \left(-\frac{q}{2} + 1, \dots, 0, 1, \dots, \left(\frac{q}{2} \right) \right) \quad (10)$$

Aus Gl. 9 und Gl. 10 folgt, dass die erste Brillouin-Zone einer SWNT aus $2N$ äquidistanten Linien besteht, die in Richtung von \vec{k}_\parallel verlaufen. Für die *zigzag*- und *armchair*-CNT, die beispielhaft in Abb. 2 (Seite 5) skizziert sind, ist die Konstruktion der ersten Brillouin-Zone hier jeweils graphisch in Abb. 5 dargestellt. Die Linien haben eine Breite von $2\pi/|\vec{T}|$ und einen Abstand von $2\pi/|\vec{C}|$, wobei die Breite nach REICH ET AL. für alle achiralen CNTs auf $2\pi/a$ bzw. $2\pi/\sqrt{3}a$ vereinfacht wird [21]. Die Indizierung der $k_{\perp,p}$ erfolgt über $p \in [-N, N]$, wobei $k_{\perp,p=0}$ durch den Γ -Punkt des Graphen-Gitters geht. Aus geometrischen Betrachtungen kann gefolgert werden, dass die Zustände für p und $-p$ und für k_\parallel bzw. $-k_\parallel$ jeweils identisch sind.

Mit Hilfe der *zone folding*-Methode kann nun das elektronische Verhalten von CNTs abgeleitet werden: Die erlaubten Zustände \vec{k} einer CNT werden durch eine eindeutig bestimmte Kurvenschar paralleler Linien beschrieben, die auf das reziproke Graphen-Gitter entlang \vec{k}_\parallel projiziert werden; weist die CNT erlaubte Zustände an den K-Punkten auf, zeigt sie (halb)metallisches Verhalten analog Graphen. Ähnlich wie für das zweidimensionalen Graphen dessen Brillouin-Zone dreidimensional dargestellt werden kann, existiert für die eindimensionalen CNTs eine zweidimensionale Darstellung. Eine solche, in Abb. 5c gezeigte Darstellung erfolgt entlang einer Linie Γ -X, wobei der X-Punkt die Grenze der Brillouin-Zone markiert. Anschaulich ausgedrückt bedeutet dies, wie in Abb. 5a und Abb. 5b skizziert, dass eine CNT metallisch ist, wenn einer der

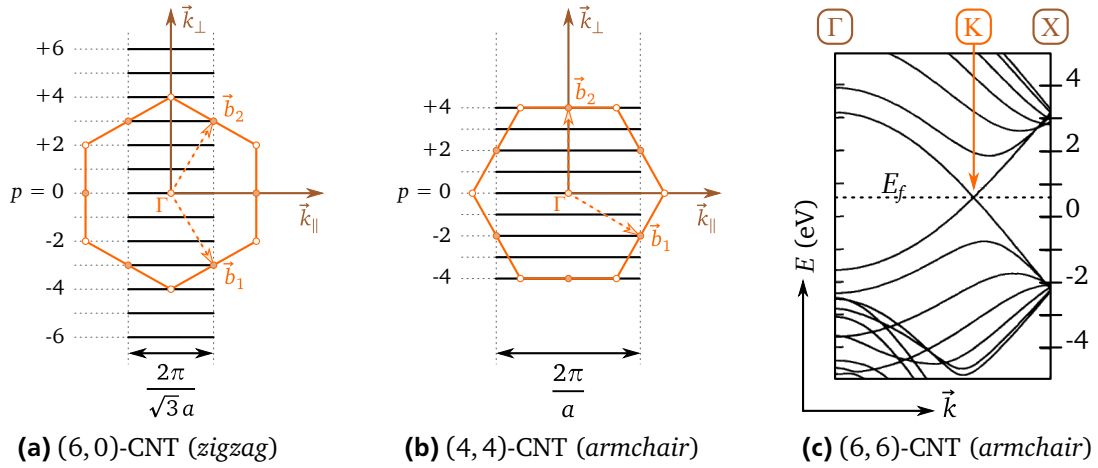


Abb. 5: (a/b) Konstruktion der ersten Brillouin-Zone der in Abb. 2a und Abb. 2b skizzierten zigzag- bzw. armchair-CNT; nach [21, 28]. Die Anzahl und Indizierung der $k_{\perp,p}$ ist über $p \in [-N, N]$ gegeben, wobei $N = \text{gcd}[n, m]$ gilt und $k_{\perp,p=0}$ durch den Γ -Punkt des Graphen-Gitters geht. (c) Brillouin-Zone entlang Γ -X, wobei der X-Punkt die Grenze der ersten Brillouin-Zone markiert; nach [20].

$k_{\perp,p}$ einen der K-Punkte des Graphen-Gitters schneidet. Unter Kenntnis der Lage der K-Punkte im Graphen-Gitter lässt sich nun nach Gl. 11 eine Abhängigkeit zwischen dem elektronischen Verhalten und dem Typ (n, m) einer CNT herstellen [21]:

$$\begin{aligned}
 \vec{K} \cdot \vec{C} &\equiv 2\pi r \quad \text{mit} \quad \vec{K} = \frac{1}{3}(\vec{b}_1 - \vec{b}_2) \\
 \Rightarrow 2\pi r &= \frac{1}{3}(\vec{b}_1 - \vec{b}_2) \cdot (n\vec{a}_1 + m\vec{a}_2) \\
 &= \frac{2\pi}{3}(n - m) \\
 \Rightarrow 3r &= n - m \quad \text{mit} \quad r \in \mathbb{Z}_0
 \end{aligned} \tag{11}$$

Erfüllt eine CNT also die Bedingung, dass die Differenz der beiden Komponenten des Chiralitätsvektors ein Vielfaches von 3 ist, so zeigt sie metallisches Verhalten. Dieser 1992 von HAMADA formulierte Zusammenhang führt unter anderem zu der Erkenntnis, dass sämtliche armchair-CNTs metallisch sind [20]. Für die halbleitenden CNTs ergibt sich in Abhängigkeit vom Durchmesser d_{SWNT} die Bandlücke E_g nach Gl. 12 [29]. Dabei ist γ_0 das Transferintegral des tight binding-Modells, das der Wechselwirkung nächster Nachbarn der $2p_z$ -Orbitale Rechnung trägt; ein typischer Wert für γ_0 ist 2,9 eV. Mit der Bindungslänge im Graphen-Gitter $a_{\text{C-C}}$ kann gemäß Gl. 13 die Bandlücke E_g einer CNT abgeschätzt werden.

$$E_g = \frac{2\pi a_{\text{C-C}} \gamma_0}{|\vec{C}|} = \frac{2a_{\text{C-C}} \gamma_0}{d_{\text{SWNT}}} \tag{12}$$

$$\Rightarrow E_g = \frac{0,8 \text{ nm}}{d_{\text{SWNT}}} \text{ eV} \quad \text{mit} \quad a_{\text{C-C}} = 1,42 \text{ \AA} \quad \text{und} \quad \gamma_0 = 2,9 \text{ eV} \tag{13}$$

Dabei sind die beiden zuerst genannten reine Verdampfungs- und Kondensationsverfahren, während es bei der CVD auch zu chemischen Reaktionen kommt. Allen gemeinsam ist der grundlegende Mechanismus, dass zunächst Kohlenstoff aus einer geeigneten Quelle in die Gasphase überführt wird und dieser anschließend auf einem Träger oder „Kollektor“ zu CNTs resublimiert. Die Resublimation kann dabei durch einen Katalysator unterstützt und beeinflusst werden, wobei die Wahl des Katalysatormaterials in einem solchen Fall eine wesentliche Rolle spielt. Oft kommen reine Katalysatoren wie Eisen, Cobalt oder Nickel zum Einsatz [30], es finden aber auch Misch-Katalysatoren wie Cobalt/Eisen [31] oder Eisen/Molybdän [32] Verwendung.

Im Folgenden sollen die drei Möglichkeiten zur Herstellung von CNTs kurz in ihren Grundzügen beschrieben werden, wobei der Aufbau der jeweiligen Anlage in Abb. 7 nach SEIDEL skizziert ist [33, Kapitel 2]:

a) Lichtbogenentladung:

Dieses Verfahren nutzt eine sich selbst erhaltende Gasentladung zwischen zwei Elektroden, wobei durch das sich bildende Plasma extrem hohe Temperaturen entstehen, die wiederum zu einer Verdampfung des Elektrodenmaterials führen. Der Aufbau besteht aus einer Unterdruckkammer, in der unter Inertgas-Atmosphäre zwei Graphit-Elektroden in einen Abstand von wenigen Millimetern gebracht werden, wobei eine der Elektroden leicht größer ist als die andere. Bei einem Strom von 100 A [34] und Temperaturen von bis zu 4000 K [35] sublimiert der Kohlenstoff der Anode und lagert sich in Form von CNTs an der Kathode ab, wobei sich zusätzlich als Nebenprodukt Fullerene und als Verunreinigung Ruß bilden. Da die Anode auf diese Weise mit der Zeit verbraucht wird, muss sie kontinuierlich nachgeführt werden, um einen konstanten Abstand zwischen den beiden Elektroden zu erhalten. Fluktuationen im Plasma müssen über den Abstand der Elektroden und die angelegte Spannung kompensiert werden.

b) Laserablation:

Der grundlegende Aufbau für die Laserablation umfasst eine Laserquelle, ein Graphit-*Target* und einen konischen Kollektor. *Target* und Kollektor befinden sich in einem Quarzrohr, wobei der Bereich um das *Target* auf 1200 °C [35] geheizt wird, während sich der Kollektor in einem ungeheizten Bereich befindet und zusätzlich wassergekühlt wird. Vor dem Aufheizen wird das Rohr zunächst evakuiert und nach Erreichen der Temperatur kontinuierlich mit einem Inertgas durchflutet, sodass um das *Target* herum ein dichter Gasstrom entsteht. Durch einen gepulsten Laserstrahl wird schließlich Kohlenstoff aus dem *Target* sublimiert, der vom Gasstrom erfasst und als Teilchendampf zum gekühlten Kollektor transportiert wird, wo er in Form von CNTs und verschiedenen Nebenprodukten resublimiert.

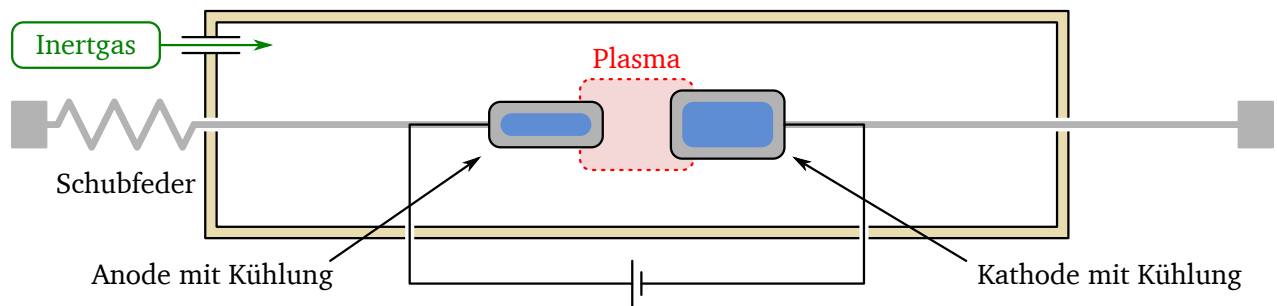
c) chemische Gasphasenabscheidung:

Zum Wachsen von CNTs mittels CVD wird eine heizbare Reaktorkammer mit einer Gaszuführung und einem Katalysator auf einem Träger benötigt. Der Katalysator, der in Form von kleinen Partikeln vorliegt, wird dabei in einem separaten Prozessschritt präpariert und gegebenenfalls auf den Träger aufgebracht. Als Katalysatormaterial haben sich Eisen, Cobalt und Nickel bewährt [36]. Bei Prozesstemperaturen zwischen 500 und 1000 °C [35] wird nach der Präparation des Katalysators ein kohlenstoffhaltiges Gas als Precursor in die Kammer eingeleitet. Die konkrete Prozesstemperatur hängt von der Wahl des Precursors ab, liegt aber unter den für die Laserablation benötigten 1200 °C. Auf der Oberfläche der Katalysatorpartikel dissoziiert der Precursor zu Kohlenstoff und Co-Produkten, die über die Gasphase mit Hilfe eines Spülgases abtransportiert werden, während der Kohlenstoff in den Partikel oder an dessen Oberfläche entlang diffundiert. Durch die begrenzte Löslichkeit von Kohlenstoff in Eisen, Cobalt oder Nickel kommt es zu einer Ausscheidung von Kohlenstoff an der Oberfläche der Partikel, was schlussendlich zum Wachstum von CNTs ausgehend von den Katalysatorpartikeln führt [37, Kapitel 3].

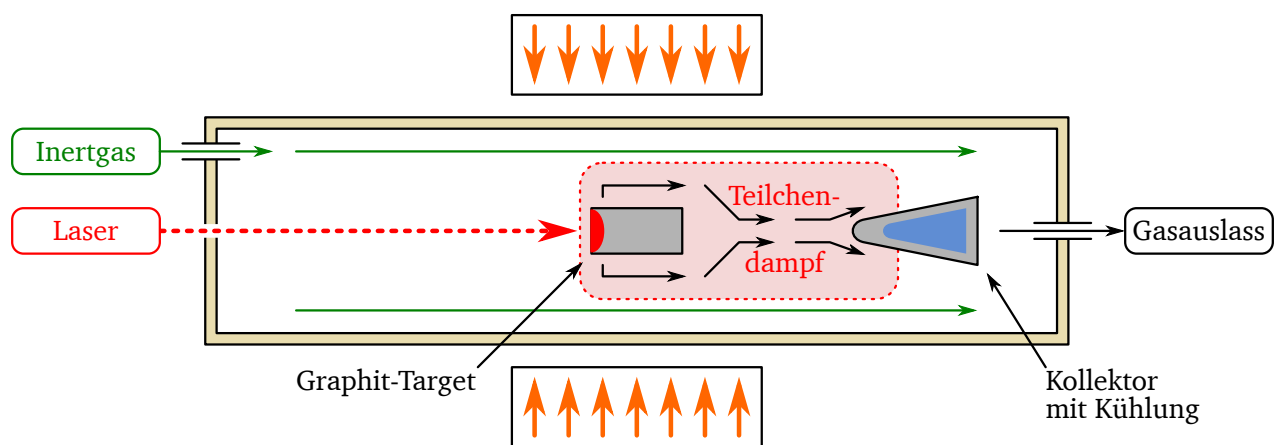
In dieser Arbeit genutztes Verfahren

Mit der vorliegenden Arbeit wird der in der Vorgängerarbeit von DR.-ING. LORRAINE RISPAL [6] entwickelte Ansatz für die Herstellung für CNTs mittels CVD weiter verfolgt. Die Details des gewählten Prozesses werden in Kapitel 7 (Seite 67) erläutert.

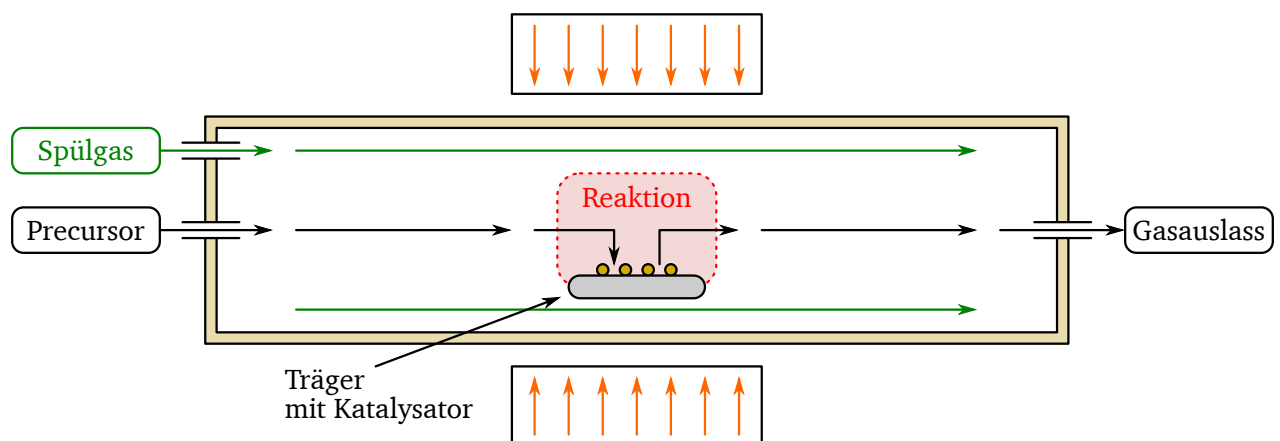
Hinsichtlich der Verwendung von CNTs in integrierten Schaltungen hat die CVD gegenüber der Lichtbogenentladung und der Laserablation den Vorteil, dass sich mit ihr ein *in situ*-Prozess realisieren lässt. Der Träger, auf dem sich der Katalysator befindet und von dem die gewachsenen CNTs nach dem Prozess sonst abgelöst werden, kann nämlich nicht nur als temporäres Hilfsmittel, sondern auch als das Substrat des Gesamtprozesses betrachtet werden. Die Lichtbogenentladung und der Laserablation dagegen sind zwingend *ex situ*-Prozesse; das bedeutet, dass das „Rohmaterial“ nach dem Prozess zuerst von der Kathode bzw. dem Kollektor extrahiert, von unerwünschten Co-Produkten wie Ruß oder Fullerenen befreit und schließlich an der gewünschten Stelle in der Schaltung positioniert werden muss. Gelingt es also, sowohl die Präparation der Katalysatorpartikel als auch den CVD-Prozess zum Wachstum der CNTs in den Gesamtprozess in geeigneter Weise zu integrieren, so ist das Entfernen, Aufbereiten und Neupositionieren der CNTs nicht mehr nötig. Die mehrfachen Transferschritte stellen nicht nur einen Mehraufwand dar, sondern sie sind auch geeignet die Struktur der CNTs zu schädigen. Mit einem *in situ*-Prozess wie der CVD, also einem Prozess, der CNTs gezielt dort wachsen lässt, wo sie im Anschluss Verwendung finden sollen, können strukturell unversehrte CNTs erzeugt werden.



(a) Lichtbogenentladung: Unter Inertgas wird ein Lichtbogen gezündet, was zu einem Plasma mit Temperaturen von bis zu 4000 K führt, die das Material der Anode verdampfen und auf der Kathode wieder kondensieren lassen. Die Anode wird mit einer Schubfeder nachgeführt, um trotz des Materialabtrags einen konstanten Abstand zur Kathode zu halten.



(b) Laserablation: Das auf 1200 °C geheizte Graphit-Target wird mit einem Laser beschossen, wodurch es verdampft. Das verdampfte Material wird durch einen dichten Strom von Inertgas zum gekühlten Kollektor transportiert, wo sich der Teilchendampf niederschlägt.



(c) CVD: Über einen auf 500 bis 1000 °C geheizten, zuvor mit Katalysatorpartikeln präparierten Träger wird ein kohlenstoffhaltiger Precursor geleitet, der dort dissoziiert. Der Kohlenstoff diffundiert in den Partikel oder an dessen Oberfläche entlang, wobei es dann zur Ausscheidung von Kohlenstoff an der Partikeloberfläche und dem Wachstum von CNTs kommt.

Abb. 7: Schematischer Aufbau der verschiedenen Anlagen für die Herstellung von CNTs nach SEIDEL [33, Kapitel 2].

3 Grundlagen zu Feldeffekt-Transistoren

3.1 Metall–Halbleiter-Kontakt

Bei der technologischen Realisierung von integrierten Schaltungen werden üblicherweise zwei Arten von Metall–Halbleiter-Kontakten unterschieden: die Schottky-Diode und der ohmsche Kontakt. Die Schottky-Diode ähnelt in ihrem gleichrichtenden Verhalten stark dem einer pn-Diode, während sich ein ohmscher Kontakt wie ein Widerstand verhält. Die prinzipiellen Effekte, die beim Zusammenführen eines Metalls und eines n-Typ-Halbleiters auftreten, sind in Abb. 8 nach PIERRET in idealisierter Weise skizziert [38, Kapitel 14]. Dabei seien φ_M und φ_S die Austrittsarbeiten des Metalls und des Halbleiters. Gilt nun $\varphi_M > \varphi_S$, d. h. die Fermi-Energie $E_{f,S}$ des Halbleiters liegt näher am Vakuumniveau E_0 als die Fermi-Energie $E_{f,M}$ des Metalls, dann bildet sich am Metall–Halbleiter-Übergang eine Raumladungszone der Weite W_{RLZ} mit einer Schottky-Barriere (SB). Die Barrierehöhe φ_{SB} wird dabei entscheidend von der verwendeten

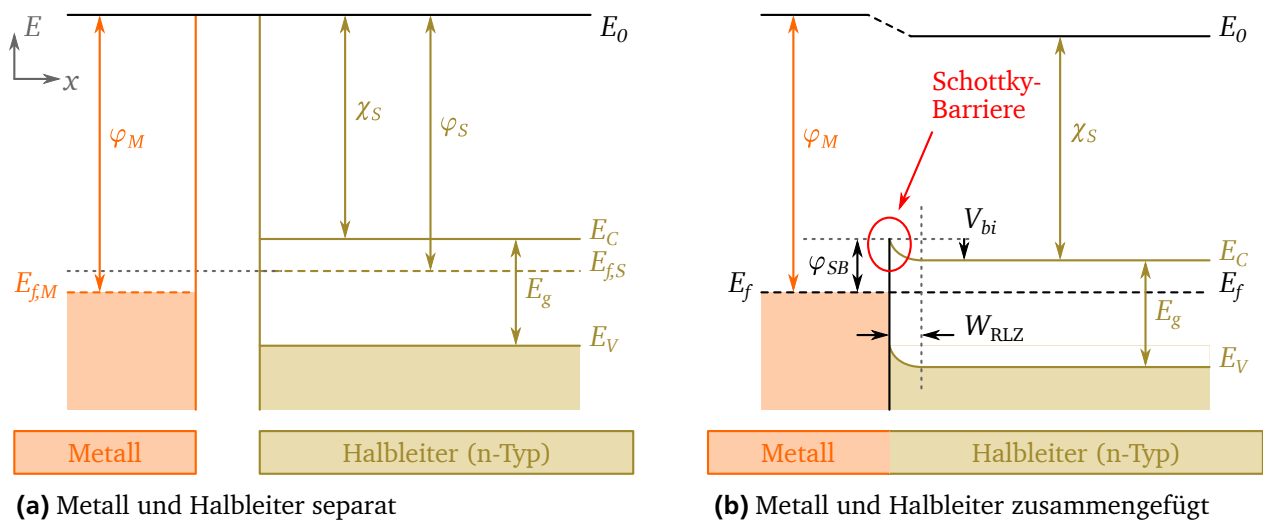


Abb. 8: Bänderdiagramm eines Metall–Halbleiter-Kontakts für ein Metall und einen n-Typ-Halbleiter nach PIERRET [38, Kapitel 14], wobei die Austrittsarbeit des Metalls φ_M größer ist als die des Halbleiters φ_S . Dabei sind: $E_{f,M}$ und $E_{f,S}$ die Fermi-Energien des Metalls und des Halbleiters; χ_S die Elektronenaffinität des Halbleiters; E_0 das Vakuumniveau; E_C und E_V die Leitungs- und Valenzbandkante und E_g die Bandlücke des Halbleiters; φ_{SB} die Höhen der Schottky-Barriere; W_{RLZ} die Weite der Raumladungszone; und V_{bi} das eingebaute Potential. (Die Elementarladung q wird der Übersicht halber unterschlagen.)

Metall–Halbleiter-Kombination beeinflusst und ist gemäß Gl. 14 über die Austrittsarbeit φ_M des Metalls und die Elektronenaffinität χ_S des Halbleiters gegeben.

$$\varphi_{SB} = \varphi_M - \chi_S \quad (14)$$

Ein Maß für die mit der Schottky-Barriere verbundenen Bandverbiegung ist das eingebaute Potential V_{bi} , das gemäß Gl. 15 über die Barrierehöhe φ_{SB} und die Valenzbandkante E_V des Halbleiters im Flachbandfall (FB) gegeben ist.

$$V_{bi} = \varphi_{SB} - \frac{(E_C - E_f)_{FB}}{q} \quad (15)$$

Wird an einen solchen Schottky-Kontakt eine Spannung V_A angelegt, kann mit einer Diodengleichung nach Gl. 16 der Strom I_{SB} , der durch die Schottky-Barriere fließt, abgeschätzt werden, wobei $I_{SB,sat}$ der Sättigungssperrstrom und V_{th} die Temperaturspannung sind.

$$I_{SB} = I_{SB,sat} \left(\exp\left(\frac{V_A}{V_{th}}\right) - 1 \right) \quad \text{mit} \quad V_{th} = \frac{k_B T}{q} \quad (16)$$

Für den hier diskutierten Fall eines n-Typ-Halbleiters mit $\varphi_M > \varphi_S$ werden die drei folgenden Spannungsbereiche für V_A unterschieden, die in Abb. 9 skizziert sind:

a) thermisches Gleichgewicht:

Ohne äußere angelegte Spannung ($V_A = 0V$) zeigt das Fermi-Niveau einen konstanten Verlauf: $E_{f,M} = E_{f,S} = E_f$. Abhängig von den Materialeigenschaften bildet sich somit eine Schottky-Barriere mit der Höhe φ_{SB} am Metall–Halbleiter-Übergang (Gl. 14), wobei auf der Seite des Halbleiters das Leitungs- und Valenzband um den Wert des eingebauten Potentials V_{bi} nach oben gebogen werden (Gl. 15).

b) Flusspolung:

Wird an die metallische Seite eine positive Spannung V_F ($V_A > 0V$) angelegt, wird $E_{f,M}$ gegenüber $E_{f,S}$ um V_F nach unten verschoben. Dadurch wird die Bandverbiegung verringert, sodass sich Majoritätsladungsträger des Halbleiters, hier Elektronen aus dem Leitungsband, zum Metall bewegen und die Schottky-Barriere überwinden können.

c) Sperrpolung:

Beim Anlegen einer negativen Spannung V_R andererseits ($V_A < 0V$) wird $E_{f,M}$ gegenüber $E_{f,S}$ um V_R nach oben verschoben. Auf diese Weise wird die Bandverbiegung weiter verstärkt und die zu überwindende Barriere vergrößert, wodurch der Elektronentransport bzw. der Stromfluss stark reduziert wird.

Für Fluss- und Sperrpolung treten dabei, abhängig von der angelegten Spannung V_A für eine gegebene Temperatur, quantenmechanische Tunnelprozesse von Ladungsträgern allein oder in Kombination mit thermischer Emission auf. Es werden die drei Transportmechanismen unterschieden, die in Abb. 9b skizziert sind, mit deren Hilfe Ladungsträger die Barriere am Metall-Halbleiter-Übergang überwinden können:

1. thermische Emission über die Barriere
2. thermionische Feldemission thermisch angeregter Ladungsträger
3. Feldemission durch die Barriere („Tunneln“)

Der Transport von Ladungsträgern durch Tunneln bzw. durch Feldemission ist dann möglich, wenn die Barriere am Metall-Halbleiter-Übergang genügend schmal ist. Im Fall der thermischen Emission besitzen Ladungsträger durch thermische Anregung so viel Energie, dass sie ohne zu tunneln die Barriere „überspringen“ können. Die Kombination aus diesen beiden Mechanismen stellt die zweistufige thermionische Emission dar: Zunächst werden Ladungsträger durch thermische Anregung auf ein höheres Energieniveau angehoben, welches aber noch unterhalb der Kante der Barriere liegt; ausgehend von diesem erhöhten Energieniveau findet im Anschluss ein Tunnelprozess durch die oben schmalere Barriere statt.

In Flusspolung ist aufgrund der Verringerung der Bandverbiegung die thermische Emission der dominierende Transportmechanismus, wobei Elektronen aus dem Leitungsband des Halbleiters in das Metall fließen. In Sperrpolung müssen wegen der umgekehrten Polarität von V_A Elektronen aus dem Metall die Schottky-Barriere in das Leitungsband des Halbleiters überwinden. Die Höhe und Breite der Barriere erfordern jedoch eine sehr hohe Ladungsträgerenergie, die einen relevanten Stromfluss verhindert. Der dennoch vorhandene Sperrstrom $I_{SB, sat}$ wird somit durch hochenergetische Elektronen des Metalls erzeugt.

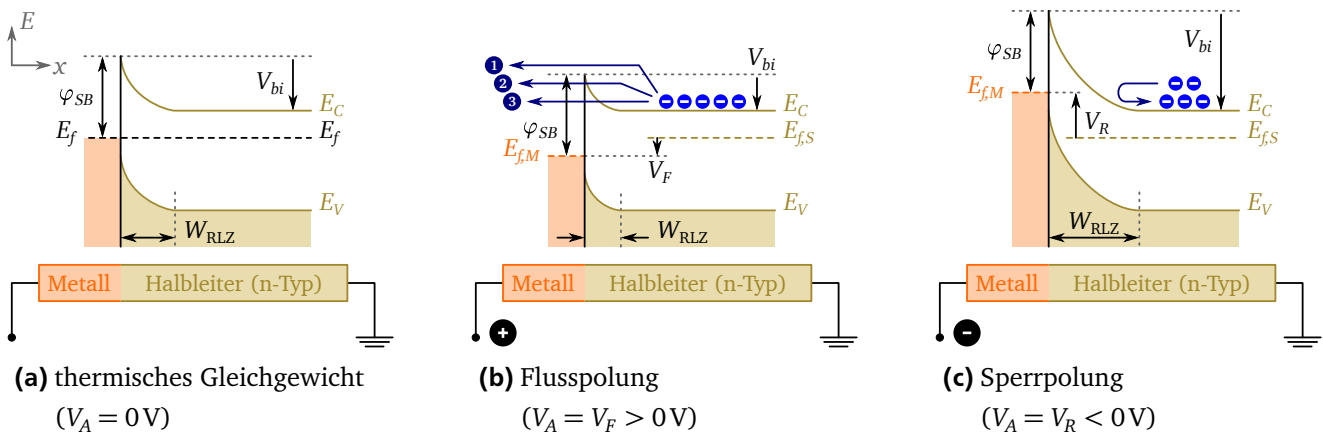


Abb. 9: Schematische Erklärung des Stromflusses durch eine Schottky-Barriere an der Grenzfläche Metall-Halbleiter nach PIERRET [38, Kapitel 14]. (Die Elementarladung q wird der Übersicht halber unterschlagen.)

3.2 MOS-Kondensator

Der Metall–Oxid–Halbleiter-Kondensator bildet als einfachste MOS-Struktur die Grundlage für den MOSFET, der im nächsten Kapitel erläutert wird. Der MOS-Kondensator besteht aus einem Halbleiter als Substrat, einem Dielektrikum sowie einem metallischen Anschluss. Der metallische Anschluss wird dabei *Gate* genannt, während der zweite Anschluss durch das Substrat gegeben ist (*Bulk*). Als Dielektrikum werden heute *high κ* -Materialien wie Hafniumdioxid verwendet, während früher meist Siliziumoxinitride eingesetzt wurden. Der MOS-Kondensator kann auf p- und auf n-dotierten Substraten ausgeführt werden und wird dabei entsprechend dem Substrat-Typ als p-MOS-Kondensator bzw. n-MOS-Kondensator bezeichnet.

Wie für den Metall–Halbleiter-Kontakt können die energetischen Zustände der Ladungsträger in einer MOS-Struktur in Form von Bänderdiagrammen dargestellt werden. Das Bänderdiagramm eines idealen n-MOS-Kondensators ist in Abb. 10 gezeigt. Nach SZE kann von einem idealen Kondensator gesprochen werden, wenn es zum einen an den Grenzflächen und im Volumen des Dielektrikums keinerlei Einfangstellen für Ladungsträger gibt und wenn zum anderen das Dielektrikum als perfekter Isolator betrachtet werden kann [39]. Das Metall und der Halbleiter sollen hier zur Vereinfachung die gleiche Austrittsarbeit besitzen, d. h. es sei $\varphi_M = \varphi_S$.

Abb. 10a zeigt den sogenannten Flachbandfall (FB), für den sämtliche Bänder einen konstanten Verlauf haben. Um diesen Zustand zu erreichen, muss im Allgemeinen eine von Null abweichende *Gate*-Spannung V_G , die sogenannte Flachbandspannung V_{FB} , angelegt werden. Aufgrund der Vereinfachung, dass die Austrittsarbeiten von Metall und Halbleiter identisch seien, gilt hier $V_{FB} = 0\text{ V}$. Wird an einen n-MOS-Kondensator eine positive Spannung $V_G > V_{FB}$ angelegt, werden zusätzliche Majoritätsladungsträger elektrostatisch zur Oxid–Substrat-Grenzfläche gezogen, sodass sich ihre Konzentration lokal erhöht. Dieser Zustand wird als Akkumulation bezeichnet und ist in Abb. 10b skizziert. Wird eine leicht negative Spannung $V_G < V_{FB}$ angelegt, werden die Bänder in die entgegengesetzte Richtung verbogen und die Majoritätsladungsträger werden von der Grenzfläche abgestoßen. Dabei bleiben die ionisierten Dotierstoffatome zurück, weswegen dieser in Abb. 10c gezeigte Zustand als Verarmung bezeichnet wird. Eine weitere Vergrößerung der (negativen) Spannung $V_G \ll V_{FB}$ führt zu einer noch stärkeren Verbiegung der Bänder nach oben; es werden nun zusätzlich Minoritätsladungsträger angezogen, sodass deren Konzentration an der Grenzfläche zum Oxid signifikant höher ist als die der Majoritätsladungsträger. Dieser Zustand wird als Inversion bezeichnet und ist in Abb. 10d skizziert.

Tab. 1 fasst den Zusammenhang zwischen der *Gate*-Spannung V_G und den verschiedenen Zuständen im Bändermodell zusammen. Hierbei wird das über V_G eingestellte Oberflächenpotential ψ_S zum *Bulk*-Potential ψ_B in Beziehung gesetzt.

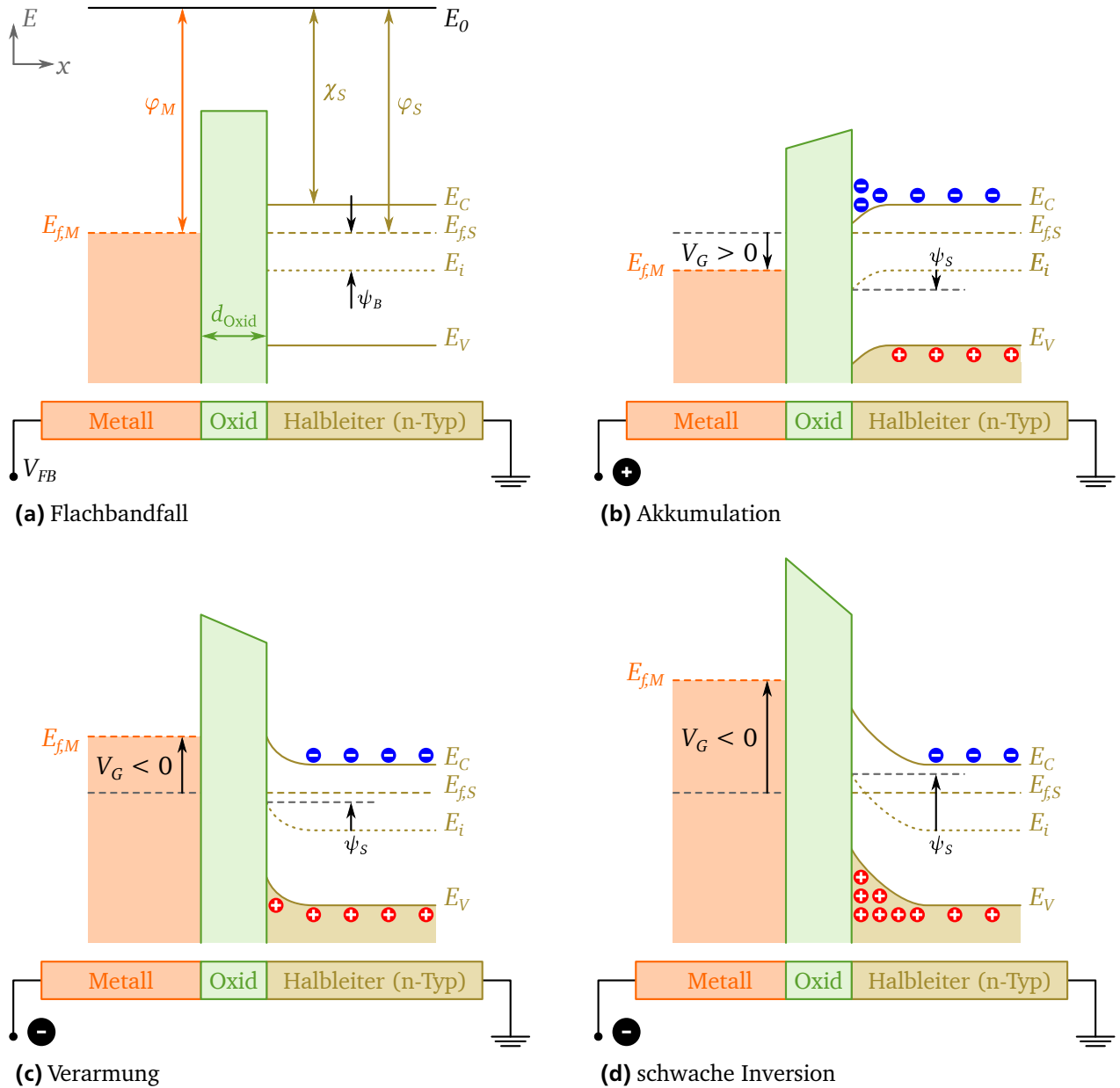


Abb. 10: Bänderdiagramm eines idealen n-MOS-Kondensators nach SZE [39, Kapitel 4]. Dabei sind: d_{Oxid} die Dicke des Oxids, ψ_B das Bulk-Potential in Bezug auf das intrinsische Energieniveau E_i und ψ_s das Oberflächenpotential. (Die Elementarladung q wird der Übersicht halber unterschlagen.)

Tab. 1: Grenzflächenzustände eines idealen n-MOS-Kondensator bezogen auf die Konzentration der Majoritätsträger nach SZE [39, Kapitel 4]. Hierbei sind ψ_B das Bulk-Potential in Bezug auf das intrinsische Energieniveau E_i und ψ_s das durch die Spannung V_G eingestellte Oberflächenpotential.

(1)	$\psi_s > 0$	Akkumulation	vgl. Abb. 10b
(2)	$\psi_s = 0$	Flachbandfall	vgl. Abb. 10a
(3)	$\psi_B < \psi_s < 0$	Verarmung	vgl. Abb. 10c
(4)	$\psi_s = \psi_B$	„intrinsischer Zustand“	
(5)	$2\psi_B < \psi_s < \psi_B$	schwache Inversion	vgl. Abb. 10d
(6)	$\psi_s < 2\psi_B$	starke Inversion	

3.3 MOS-Feldeffekt-Transistor

In seiner Funktion als spannungsgesteuerter Schalter hat der Metall–Oxid–Halbleiter-Feldeffekt-Transistor (MOSFET) seinerzeit die Elektronenröhren abgelöst und stellt nun die wichtigste Komponente in digitalen elektronischen Schaltkreisen dar. Im Folgenden sollen knapp das Funktionsprinzip und die für diese Arbeit relevanten Begriffe und Größen eines Feldeffekt-Transistors erläutert werden.

Der Zusammenhang zwischen Bezeichnung und Substrat-Typ verhält sich beim MOSFET umgekehrt wie beim MOS-Kondensator: der p-Kanal-MOSFET (p-MOSFET) hat ein n-dotiertes und der n-Kanal-MOSFET (n-MOSFET) ein p-dotiertes Substrat. In der Literatur wird in den meisten Fällen für die Beschreibung des MOSFET der n-MOSFET als Anschauungsobjekt gewählt. Da die in dieser Arbeit hergestellten Transistoren aber p-MOS-Verhalten zeigen, soll hier davon abweichend der p-MOSFET betrachtet werden, wobei die dargestellten Zusammenhänge qualitativ aber auch für den n-MOSFET gelten.

Wie Abb. 11a zeigt, hat der MOSFET im Vergleich zum MOS-Kondensator zwei weitere Anschlüsse. Zusätzlich zum *Gate*- und *Bulk*-Anschluss und dem dazwischen befindlichen Dielektrikum besitzt er einen *Drain*- und *Source*-Anschluss. Der *Drain*- und der *Source*-Anschluss werden dabei in Form von stark p-dotierten Bereichen realisiert, die in das n-dotierte Substrat eingelassen sind und von oben über eine Metallisierung kontaktiert werden. Auf diese Weise entsteht unterhalb des Dielektrikums eine pnp-Struktur, die zunächst keinen Stromfluss zulässt, da sich aufgrund des wechselnden Typs der Majoritätsladungsträger an den Grenzflächen Raumladungszonen bilden. Beim p-MOSFET werden der *Source*- und *Bulk*-Anschluss kurzgeschlossen und die *Source/Drain*-Spannung V_{DS} als negatives Potential an den *Drain*-Kontakt angelegt.

Abb. 11b skizziert die drei Arbeitsbereiche eines MOSFET: Sperr-, Trioden- und Sättigungsbereich. Wird eine Spannung V_{GS} an den *Gate*-Kontakt angelegt, die größer als die Schwellspannung V_T des Transistors ist ($V_{GS} > V_T$, vgl. Abb. 11c), so befindet sich der *Gate*-Kondensator im Zustand der Verarmung und es besteht zwischen *Source* und *Drain* kein leitfähiger Kanal — der Transistor befindet sich im Sperrbereich. Wird dagegen ein $V_{GS} < V_T$ angelegt, bildet sich zwischen *Source* und *Drain* ein Inversionskanal in Form von Löchern bzw. Defektelektronen und es kann für kleine Werte der *Source/Drain*-Spannung V_{DS} ein *Source/Drain*-Strom I_{DS} fließen, der zunächst linear und anschließend nach einer Wurzel-Funktion mit V_{DS} skaliert — der Transistor befindet sich im Triodenbereich. Im Triodenbereich kann der Strom I_{DS} mit Gl. 17 beschrieben werden [38, Kapitel 17]. Hierbei sind W und L die Weite bzw. die Länge des Kanals, μ die Ladungsträgerbeweglichkeit und C die Kapazität des *Gate*-Kondensators.

$$I_{DS} = \frac{W}{L} \mu C \left((V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right) \quad \text{mit} \quad C = \varepsilon_0 \left(\frac{\varepsilon_r}{d} \right)_{\text{Dielektrikum}} \quad (17)$$

Erreicht V_{DS} die Sättigungsspannung $V_{DS, sat}$, wird keine weitere Erhöhung von I_{DS} erzielt, sondern der Strom bleibt unabhängig von V_{DS} konstant — der Transistor befindet sich im Sättigungsbereich. Für eine gegebene *Gate*-Spannung V_{GS} ist $V_{DS, sat}$ über Gl. 18 definiert, während der Sättigungsstrom $I_{DS, sat}$ über Gl. 19 bestimmt wird.

$$V_{DS, sat} \equiv V_{GS} - V_T \quad (18)$$

$$I_{DS, sat} = \frac{W}{L} \mu C \frac{(V_{GS} - V_T)^2}{2} \quad (19)$$

Abb. 11c und Abb. 11d zeigen schematisch die Eingangskennlinie bzw. die Unterschwellenkennlinie eines p-MOSFET. Es wird dabei der *Source/Drain*-Strom I_{DS} für eine gegebene *Source/Drain*-Spannung V_{DS} in Abhängigkeit von der *Gate*-Spannung V_{GS} linear bzw. halblogarithmisch dargestellt.

Aus einer gemessenen Eingangskennlinie kann die Schwellspannung V_T des Transistors nach SCHRODER mit der Methode der sogenannten „linearen Extrapolation“ graphisch bestimmt werden [40, Kapitel 4]. Es wird dazu der lineare Bereich der Kennlinie verlängert und der Schnittpunkt $V_{T, extrapoliert}$ mit der V_{GS} -Achse bestimmt. Mit Hilfe des Schnittpunkts und Gl. 17 ist die Schwellspannung V_T dann über Gl. 20 gegeben.

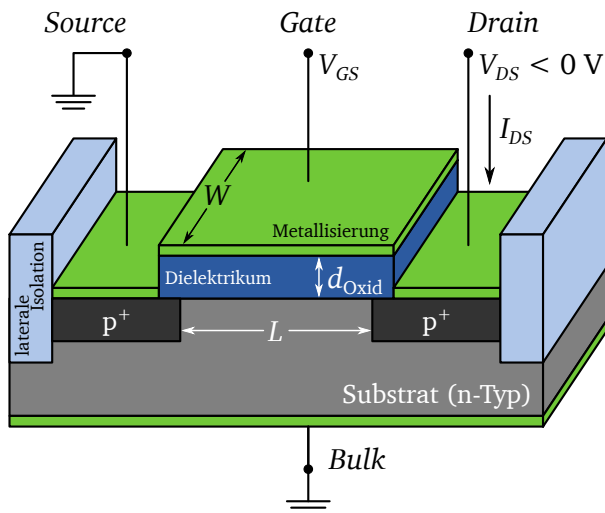
$$V_T = V_{T, extrapoliert} - \frac{V_{DS}}{2} \quad (20)$$

Die Unterschwellenkennlinie zeigt zwei Bereiche, in denen I_{DS} durch die halblogarithmische Darstellung als beinahe konstant erscheint; sie werden als *On*- und *Off*-Zustand und die entsprechenden Ströme als *On*- und *Off*-Strom (I_{on} bzw. I_{off}) bezeichnet. Als Maß für den Unterschied zwischen I_{on} und I_{off} wird das *On/Off*-Verhältnis angegeben, das gemäß Gl. 21 als Quotient der Ströme definiert ist.

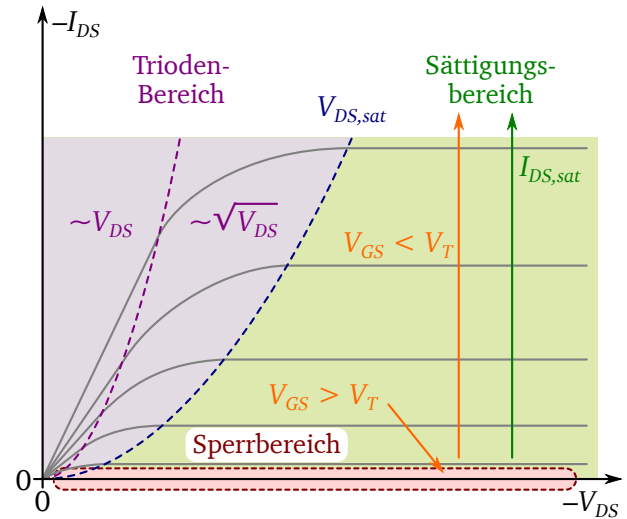
$$On/Off \equiv \frac{I_{on}}{I_{off}} \quad (21)$$

Schließlich kann die Unterschwellensteigung S_{subthr} des Transistors ermittelt werden. Sie ist per Definition nach SZE gemäß Gl. 22 gegeben und stellt ein Maß dafür dar, wie abrupt der *Source/Drain*-Strom I_{DS} abfällt bzw. ansteigt [39, Kapitel 6]. Die Unterschwellensteigung kann graphisch im Übergangsbereich zwischen *On*- und *Off*-Zustand als Steigung der Kennlinie ermittelt werden.

$$S_{subthr} \equiv \frac{\Delta V_{GS}}{\text{Dekade } I_{DS}} \quad \text{wobei} \quad [S_{subthr}] = 1 \text{ mV/dec} \quad (22)$$

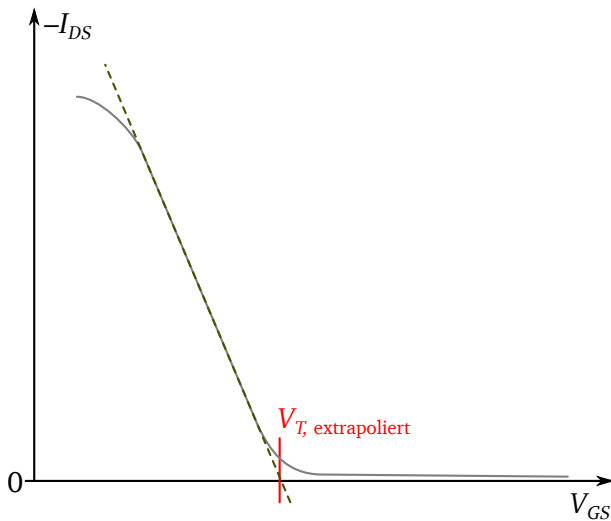


(a) schematische Darstellung



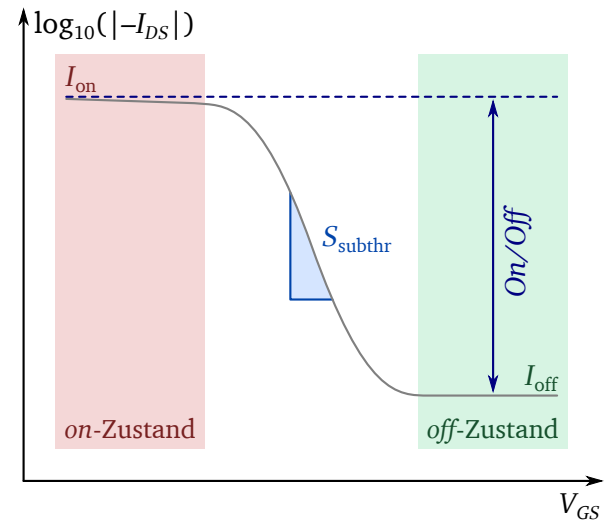
(b) Ausgangskennlinie

(I_{DS} vs. V_{DS} für verschiedene V_{GS})



(c) Eingangskennlinie

(I_{DS} vs. V_{GS} , lineare Darstellung)



(d) Unterschwellenkennlinie

(I_{DS} vs. V_{GS} , halblogarithmische Darstellung)

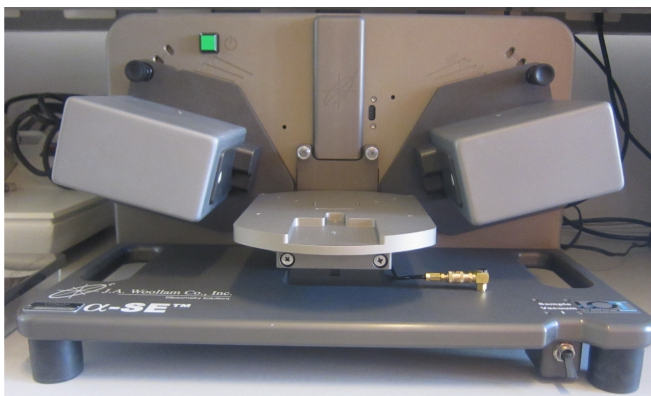
Abb. 11: Schematische Darstellung und idealisierte Kennlinien eines p-Kanal-MOSFET nach SZE [39, Kapitel 6] und SCHRODER [40, Kapitel 4]. Dabei sind: W und L die Weite bzw. die Länge des Kanals und d_{Oxid} die Dicke des Gate-Dielektrikums; V_{GS} und V_{DS} die Gate- bzw. die Source/Drain-Spannung; I_{DS} der Source/Drain-Strom; $V_{DS, \text{sat}}$ und $I_{DS, \text{sat}}$ die Sättigungsspannung bzw. der Sättigungsstrom. (c/d): Als Kenngrößen des Transistors können aus gemessenen Kennlinien die Schwellenspannung V_T und die Unterschwellensteigung S_{subthr} graphisch extrahiert werden (Gl. 20 bzw. Gl. 22), während das On/Off-Verhältnis des Source/Drain-Stroms als Quotient von I_{on} und I_{off} bestimmt wird (Gl. 21).

4 Analytische Methoden

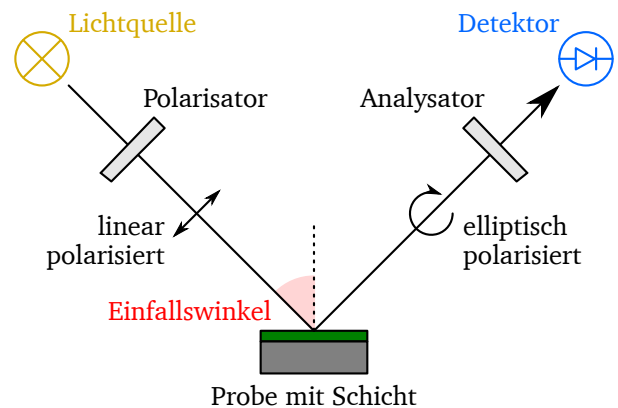
4.1 Ellipsometrie

Die experimentelle Bestimmung der Schichtdicken der gewachsenen Oxide erfolgt in dieser Arbeit mit Hilfe des in Abb. 12a gezeigten spektroskopischen Ellipsometers *alpha-SE* des Herstellers *Woollam*. Bei einer ellipsometrischen Messung wird die Polarisationsänderung von Licht ausgenutzt, die bei Reflexion an einer Oberfläche und Transmission einer Schicht auftritt. Wie stark hierbei die Polarisationsänderung ausfällt, hängt von den optischen Eigenschaften und der Schichtdicke der entsprechenden Materialien ab. Ein Ellipsometer besteht wie in Abb. 12b skizziert im Wesentlichen aus einer Lichtquelle, einem Polarisator, einem Analysator und einem Detektor.

Das zunächst unpolarisiert Licht der Quelle wird durch den Polarisator linear polarisiert, d. h. die Vektoren der elektromagnetischen Welle werden bei konstantem Betrag auf jeweils eine einzige Richtung senkrecht zur Ausbreitungsrichtung beschränkt. Durch die Wechselwirkung mit der Probe ändert sich der Polarisationszustand von linear zu elliptisch, d. h. die Vektoren der elektromagnetischen Welle rotieren mit veränderlichem Betrag senkrecht um die Ausbreitungsrichtung. Das elliptisch polarisierte Licht gelangt durch einen Analysator in den Detektor, wo es in ein elektrisches Signal umgesetzt wird. Die Intensität des am Detektor ankommenden Lichts variiert dabei je nach Winkelerorientierung des Analysators. Aus den daraus erhaltenen Informationen können die ellipsometrischen Parameter Ψ und Δ bestimmt werden. Über die



(a) Ellipsometer *alpha-SE* des Herstellers *Woollam*



(b) Schematischer Aufbau

Abb. 12: Zur Bestimmung der Schichtdicken der verwendeten Dielektrika genutztes Ellipsometer am IHTN.

ellipsometrische Grundgleichung gemäß Gl. 23 hängen diese Parameter mit den komplexen Reflexionskoeffizienten R^{\parallel} und R^{\perp} für parallel bzw. senkrecht zur Einfallsebene polarisiertem Licht zusammen.

$$\tan(\Psi) \exp(i\Delta) = \frac{R^{\parallel}}{R^{\perp}} \quad (23)$$

Wird von einer planparallelen, ideal glatten Schicht ausgegangen, so lassen sich die Reflexionskoeffizienten R^{\parallel} und R^{\perp} mit Hilfe des in Abb. 13 skizzierten Einschichtmodells berechnen. Der unter einem bestimmten Winkel φ_0 einfallende Lichtstrahl wird an den Grenzflächen Luft–Schicht und Schicht–Substrat unterschiedlich gebrochen und reflektiert. Dabei gelten für die Reflexion und Transmission des Lichts an den jeweiligen Grenzflächen die Fresnel’schen Gleichungen nach Gl. 24 bis Gl. 27 für parallel bzw. senkrecht polarisiertes Licht.

$$r_{ij}^{\parallel} = \frac{n_j \cos(\varphi_i) - n_i \cos(\varphi_j)}{n_j \cos(\varphi_i) + n_i \cos(\varphi_j)} \quad (24)$$

$$r_{ij}^{\perp} = \frac{n_i \cos(\varphi_i) - n_j \cos(\varphi_j)}{n_i \cos(\varphi_i) + n_j \cos(\varphi_j)} \quad (25)$$

$$t_{ij}^{\parallel} = \frac{2n_i \cos(\varphi_i)}{n_i \cos(\varphi_j) + n_j \cos(\varphi_i)} \quad (26)$$

$$t_{ij}^{\perp} = \frac{2n_i \cos(\varphi_i)}{n_i \cos(\varphi_i) + n_j \cos(\varphi_j)} \quad (27)$$

Das gesamte reflektierte Licht ergibt sich aufgrund der Mehrfachreflexion innerhalb der Schicht durch die kohärente Überlagerung der reflektierten Teilstrahlen. Mit Hilfe einer Reihenentwicklung ergeben sich die Reflexionskoeffizienten R^{\parallel} und R^{\perp} des parallel und senkrecht polarisierten Lichts schließlich gemäß Gl. 28 und Gl. 29, wobei β die Phasenverschiebung zwischen zwei be-

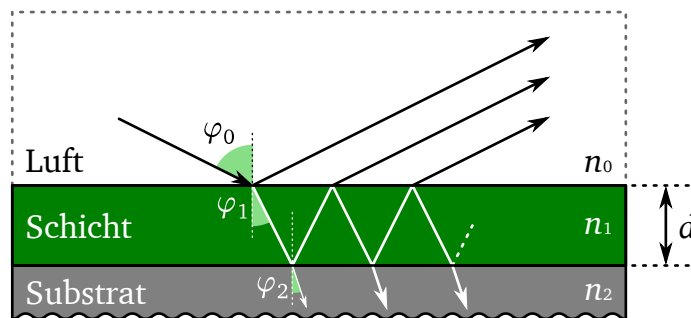


Abb. 13: Einschichtmodell der Ellipsometrie nach FUJIWARA für ein Substrat mit einer dielektrischen Schicht [41, Kapitel 1]: Der einfallende Lichtstrahl wird in Abhängigkeit von den verschiedenen Brechungsindizes n_i an den Grenzflächen Luft–Schicht und Schicht–Substrat unterschiedlich gebrochen.

nachbarten, reflektierten Teilstrahlen in Abhängigkeit von der Schichtdicke d gemäß Gl. 30 ist.

$$R^{\parallel} = \frac{r_{01}^{\parallel} + r_{12}^{\parallel} \exp(-i 2\beta)}{1 + r_{01}^{\parallel} r_{12}^{\parallel} \exp(-i 2\beta)} \quad (28)$$

$$R^{\perp} = \frac{r_{01}^{\perp} + r_{12}^{\perp} \exp(-i 2\beta)}{1 + r_{01}^{\perp} r_{12}^{\perp} \exp(-i 2\beta)} \quad (29)$$

$$\beta = 2\pi \frac{dn_1}{\lambda} \cos(\varphi_1) = 2\pi \frac{d}{\lambda} \sqrt{n_1^2 - n_0^2 \sin^2(\varphi_0)} \quad (30)$$

Da die Phasenverschiebung β direkt von der Schichtdicke d abhängt, ist über die ellipsometrische Grundgleichung ein direkter Zusammenhang zwischen der Schichtdicke und dem experimentell bestimmbareren Werten Δ bzw. Ψ gegeben. Die Phasenverschiebung zwischen zwei benachbarten, reflektierten Teilstrahlen hängt jedoch noch zusätzlich vom Brechungsindex der Schicht n_1 ab. Dadurch gibt es mehrere Kombinationen von d und n_1 , die mit den gemessenen Δ und Ψ korrespondieren können. Außerdem müssen Abweichungen beachtet werden, die z. B. durch die Rauheit der Oberfläche hervorgerufen werden können. Die Bestimmung der Schichtdicke muss somit wie in Abb. 14 skizziert iterativ erfolgen: Nachdem die gemessenen Daten gesammelt wurden, wird ein Modell konstruiert, welches dem zu erwartenden Ergebnis möglichst nahe kommt. Dieses Modell liefert mit Hilfe der Fresnel'schen Gleichungen eine Vorhersage über die Messergebnisse. Die kalkulierten Werte werden dann mit den experimentellen Werten verglichen. Schließlich werden die Parameter des Modells so lange optimiert, bis die beste Übereinstimmung zwischen dem aus dem Modell ermittelten Daten und den experimentellen Daten gefunden wurde. Daraus folgt dann schließlich die Schichtdicke d und der Brechungsindex der Schicht n_1 .

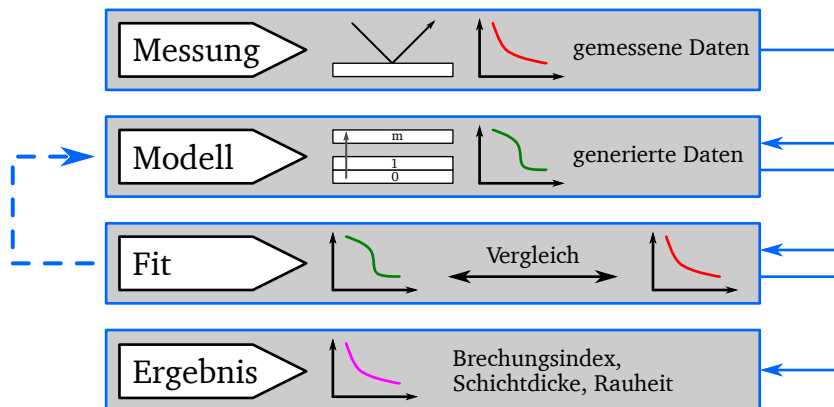


Abb. 14: Iteratives Verfahren zur Auswertung ellipsometrischer Messungen zur Bestimmung der Schichtdicke d und des Brechungsindex n .

4.2 Rasterkraftmikroskopie

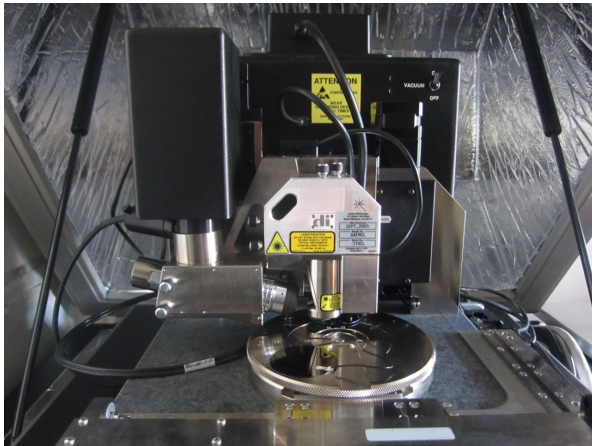
Die Rasterkraftmikroskopie (AFM) wird in dieser Arbeit zum einen zur Kontrolle der Schichtdicke nach der Abscheidung des Katalysators eingesetzt, zum anderen wird mit ihr nach der Cluster-Bildung und dem CNT-Wachstum das Prozessergebnis untersucht. Das hier genutzte Rasterkraftmikroskop ist das in Abb. 15a abgebildete *Dimension 3100* des Herstellers *Digital Instruments / Veeco*. Die Ansteuerung des Mikroskops erfolgt über die Software *NanoScope* des Herstellers in Version 6.12r1, während für die Auswertung die freie Software *Gwyddion* [42] verwendet wird.

Wie auch in Abb. 15b skizziert, besteht das Grundprinzip der AFM in einer über eine Probenoberfläche geführten Spitze (*tip*), auf welche verschiedene Kräfte durch die Proben einwirken. Die Spitze befindet sich an einem dünnen, blattfederartigen Ausleger (*cantilever*), welcher in einer Halterung fixiert ist. Wird ein Laser auf den *cantilever* gerichtet und dessen Reflexion mittels einer Photodiode vermessen, so lassen sich kleinste Bewegungen der Spitze detektieren. Der Laser agiert dabei als Verstärker, da bereits geringste Winkeländerungen zu einer nachweisbaren Verschiebung des Reflexionspunktes führen. Um eine genaue Positionierung der Spitze zu ermöglichen, wird diese mittels mehrerer piezoelektrischer Aktoren (*x**y*-Achsen) bewegt. Dies ermöglicht laterale Auflösungen im Nanometerbereich [43]. Höhenunterschiede der Probe werden mittels eines weiteren piezoelektrischen Aktors (*z*-Achse) ausgegeregelt. (Eine Ausnahme stellen Messungen mit konstanter Höhe dar, welche jedoch nur bei nahezu topologiefreien Proben angewendet werden können.) Die Photodiode besteht aus zwei oder auch vier aneinander angrenzenden Elementen, auf deren Mitte die Reflexion des Lasers gerichtet wird. Eine Ablenkung erzeugt so ein höheres Signal in einem der Elemente und ein niedrigeres Signal in einem anderen Element der Diode. Durch Auswertung dieser Differenz ist es möglich, ein besseres Signal-Rausch-Verhältnis zu erhalten als es mit einer einzelnen Diode möglich wäre.

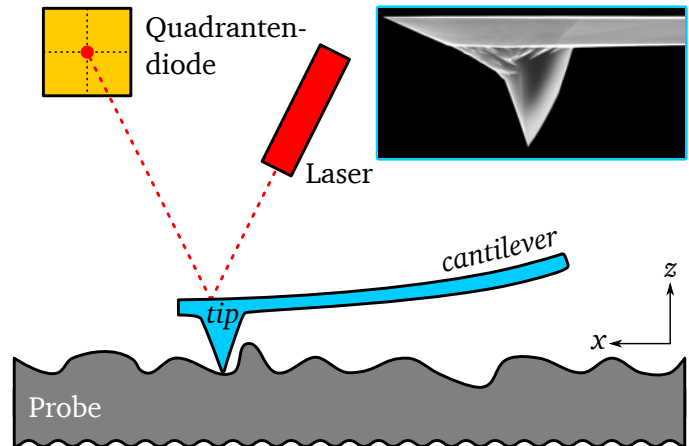
Messmethoden

AFM-Messungen können mit verschiedenen Messmethoden durchgeführt werden, welche sich in drei grundsätzliche Modi einteilen lassen: Kontakt-Modus (*contact mode*), Nicht-Kontakt-Modus (*non contact mode*) und intermittierender Modus (*tapping mode*).

Im **contact mode** wird die Messspitze so nah an die Probe gebracht, dass ein dauerhafter mechanischer Kontakt auftritt. Es handelt sich hierbei um den einfachsten und auch ältesten Modus, welcher bereits im ersten AFM von 1986 Verwendung fand [45]. Der *cantilever* wird dabei bis zu einer vordefinierten Verbiegung auf die Probe gedrückt. Diese dient als Nullpunkt, bei welcher der Laser in die Mitte der Photodioden fällt. Ändert sich die Topologie der Probe, so ändert sich aufgrund der veränderten Kräfte auch die Verbiegung, welche eine Verschiebung der Reflexion hervorruft [43]. Wird mit konstanter Höhe gemessen, so findet eine direkte Auswertung



(a) AFM Dimension 3100 des Herstellers
Digital Instruments / Veeco



(b) Schematische Darstellung des Messkopfes
(SEM-Bild der Spitze aus [44])

Abb. 15: Zur Untersuchung von Cluster-Bildung und CNT-Wachstums genutztes Rasterkraftmikroskop am IHTN.

dieses Signals statt. Alternativ — und bei Proben mit stärkeren Höhenänderungen üblich — wird die Verbiegung mittels einer Regelschleife über den z -Achsen-Aktor konstant gehalten. Die jeweils benötigte Nachregelung dient in diesem Verfahren als Messwert. Wird eine Quadrantendiode für die Detektion des Lasers verwendet, so kann auch die Verkipfung des *cantilevers* gemessen werden, wodurch sich Rückschlüsse auf die Reibung zwischen der Spitze und der Probe ziehen lassen. Wird zwischen einer leitfähigen Spitze und der Probe eine Spannung angelegt, so kann mittels eines empfindlichen Verstärkers der resultierende Strom gemessen werden. Dies ermöglicht eine orts aufgelöste elektrische Charakterisierung der Probe. Dies wird auch als Strom-Spannungs-Mikroskopie (CS-AFM, *current sensing atomic force microscopy*) bezeichnet.

Bei Messungen im **non contact mode** wird die Spitze oberhalb der Probenoberfläche gehalten, ohne dass eine Berührung eintritt. Die Spitze wird nun mit einer Schwingung angeregt. Dies wird üblicherweise mit einem weiteren piezoelektrischen Aktor realisiert. Dabei kann zwischen konstanter Frequenzanregung und Resonanzanregung unterschieden werden. Nähert sich die Spitze der Probenoberfläche, so findet durch die auftretenden Oberflächenkräfte eine Dämpfung der Amplitude statt und die Resonanzfrequenz verschiebt sich [46, Kapitel 3]. Ein Regelkreis sorgt in beiden Anregungsfällen durch Variation des Abstands für eine konstante Amplitude bzw. Resonanzfrequenz. Das Messsignal ergibt sich aus der Regelgröße. Durch Veränderung der Spitze ist es möglich, Magnetfelder oder auch chemische Eigenschaften einer Probe zu vermessen. Eine weitere Methode stellt die Raster-Kelvin-Mikroskopie (KPFM, *kelvin probe force microscopy*) dar. Zusätzlich zu einer mechanischen Anregung wird dabei eine sinusförmige Spannung mit einer überlagerten Gleichspannung an eine leitfähige Spitze angelegt. Die Gleichspannung wird dann kontinuierlich während der Messung variiert, sodass die auftretenden elektrostatischen Kräfte minimal werden, wobei die benötigte Spannung den Messwert darstellt. So lassen sich lokale Ladungsbereiche und Austrittsarbeiten bestimmen [47, Kapitel 4].

Der **tapping mode** ist eine Erweiterung der vorangegangenen Messmodi und der Modus, der in dieser Arbeit hauptsächlich verwendet wird. Er verbindet den Vorteil der geringen Interaktion von Spitze und Probe des *non contact mode* mit dem Vorteil der direkten Messwerte des *contact mode*. Wie im *non contact mode* wird eine Spitzenschwingung im Bereich der Resonanzfrequenz mittels eines piezoelektrischen Aktors erzeugt. Die Spitze wird jedoch so nah an die Probe gebracht, dass eine kurze Berührung stattfindet. Die Amplitude muss groß genug gewählt sein, damit ein Wiederloslösen noch möglich ist, wobei ein Regelkreis die Amplitude durch Variation des Abstands konstant hält [43]. Über die Regelgröße („Amplituden-Signal“) lassen sich hierbei die Messwerte („Höhen-Signal“) berechnen. Das Höhen-Signal repräsentiert dabei in direkter Weise die Topologie der Probenoberfläche. Aufgrund der weiter unten diskutierten Einschränkungen bei AFM-Messungen hat das Höhen-Signal allerdings einen schlechten Kantenkontrast. Eine gute Abbildung von Höhenänderungen, d. h. Kanten, zeigt dagegen das Amplituden-Signal, aus welchem allerdings keine direkten Höheninformationen gewonnen werden können. Ein Vergleich von Höhen- und Amplituden-Signal ist in Abb. 16 an einer exemplarischen Messung gezeigt.

Messartefakte und Einschränkungen der Auflösung

Eine AFM-Messung kommt, mathematisch ausgedrückt, durch die Faltung der Spitzengeometrie mit der Probenoberfläche zustande. Dabei hängt die Genauigkeit der gewonnenen Daten vom Größenverhältnis von Spitze zu Probenoberfläche ab: je größer die Spitze im Vergleich zu den Strukturen auf der Oberfläche ist, desto ungenauer werden diese abgebildet. Besondere Bedeutung bei der Beurteilung von AFM-Aufnahmen kommt deswegen der Spitzengeometrie zu, da mit ihr häufig Messartefakte und die damit einhergehenden Einschränkungen erklärt werden können.

Hochauflösende präzise Messergebnisse mittels AFM zu erhalten, ist aufgrund verschiedener Faktoren eine anspruchsvolle Aufgabe. Zunächst muss das verwendete Gerät selbst natürlich in

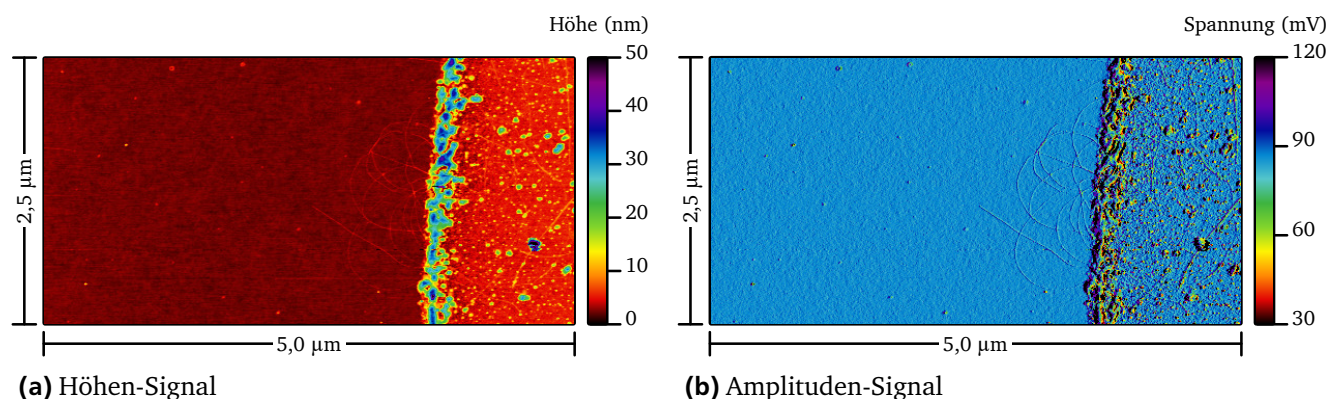


Abb. 16: AFM-Messung im *tapping mode*: Das Höhen-Signal liefert direkt Information über die Probenoberfläche, während das Amplituden-Signal Höhenänderungen, d. h. Kanten, kontrastreich zeigt.

der Lage sein, die gewünschten Strukturen abzubilden. Ebenso wichtig wie die Leistungsfähigkeit des Geräts ist allerdings auch sein Schutz gegen störende Einflüsse wie Schall und Erschütterungen. Schließlich ist auch ein gewisse Erfahrung des Anwenders nötig, da durch falsche Wahl der Messparameter zusätzliche Messartefakte erzeugt werden, die die Auswertung weiter erschweren. Gute Mikroskope erreichen eine laterale Auflösung im sub-Nanometerbereich und eine vertikale Auflösung im sub-Angströmbereich [48], wobei die tatsächliche Auflösung aber immer von der Spitzengeometrie und der Rauheit der Probenoberfläche abhängt.

Abb. 17a zeigt eine schematische Darstellung einer Messspitze. Die charakteristischen Kenngrößen einer Spitze sind ihr Krümmungsradius und ihr Flankenwinkel. Mit Hilfe dieser beiden Größen lässt sich das Zustandekommen der in Abb. 17b skizzierten Abweichungen zwischen dem Messsignal und der tatsächlichen Probenoberfläche nach STEFANOV anschaulich erklären [48]:

1. Hohe Strukturen erscheinen verbreitert: Die Verbreiterung am höchsten Punkt der Struktur entspricht dem Spitzenradius und nimmt entlang der Flanken nach unten hin zu. Je schmaler eine Struktur ist, desto ausgeprägter ist dieser Effekt.
2. Scharfe Kanten erscheinen rund: Sowohl spitze als auch stumpfe Winkel werden aufgrund der nicht punktförmigen Spitze und ihrer Krümmung abgerundet, wobei auch hier der Effekt für kleine Strukturen am ausgeprägtesten ist.
3. Begrenzte Steigung der Flanken: Nur Flanken, deren Steigung geringer ist als der Flankenwinkel der Spitze, werden korrekt abgebildet, während für steilere Flanken immer der Flankenwinkel der Spitze dargestellt wird.

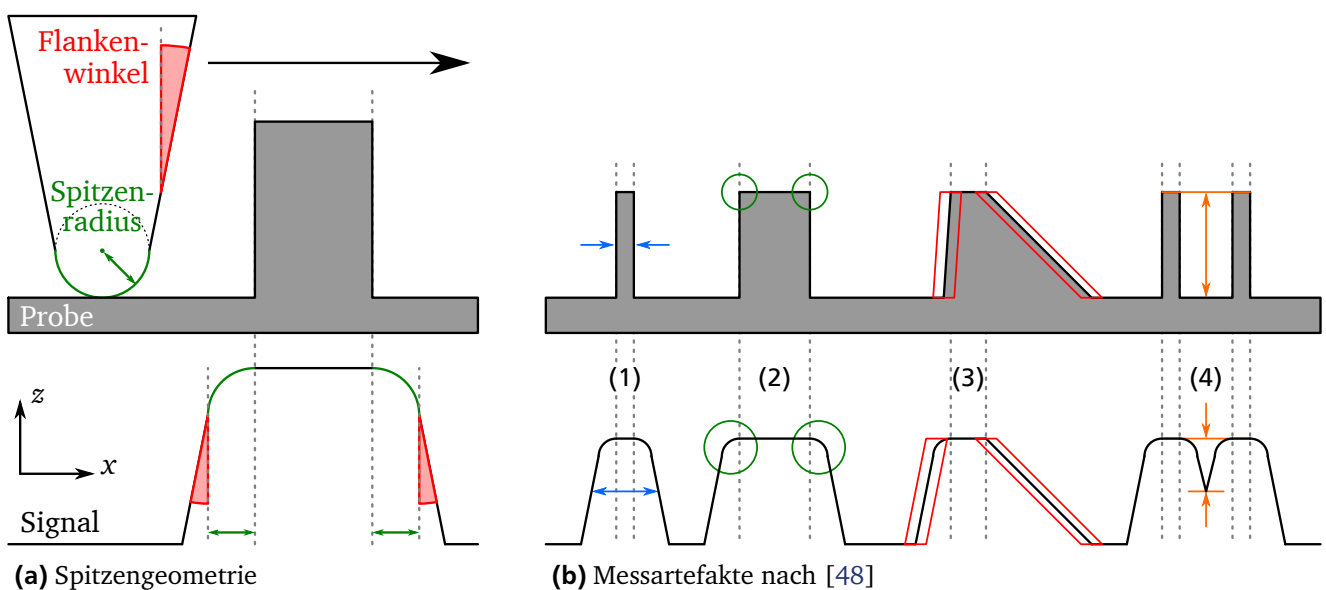


Abb. 17: Abweichungen zwischen Messsignal und der tatsächlichen Topologie der Probenoberfläche bei AFM-Messungen: Durch die nicht punktförmige Spitze, die einen Radius und einen Flankenwinkel hat, werden Kanten verrundet und Strukturen verbreitert dargestellt.

4. Die Breite und Tiefe von Gräben erscheinen reduziert: Analog zur Verbreiterung von hohen Strukturen werden Gräben im unteren Bereich schmaler abgebildet. Für sehr tiefe und/oder sehr schmale Gräben wird unter Umständen nur eine reduzierte Tiefe dargestellt, falls die Spitze ab einer gewissen Höhe breiter als der Graben ist.

4.3 Elektrische Charakterisierung

Die elektrische Charakterisierung der in dieser Arbeit prozessierten Bauelemente erfolgt mit dem in Abb. 18 gezeigten Messaufbau, der aus dem halbautomatischen Waferprober 681 A und dem Messsystem SCS-4200 der Hersteller *Rucker and Kolls* bzw. *Keithley* besteht. Mit diesem Messaufbau werden computergestützt Strom-Spannungs-Kennlinien aufgenommen. Ein wichtiger Bestandteil dieses Messaufbaus sind die SMUs (*source measure unit*), die sowohl als Strom- oder Spannungsquelle dienen, gleichzeitig aber auch als Volt- und Amperemeter fungieren. Sie bilden dabei die Schnittstelle zwischen dem Computer, über welchen die Messparameter festgelegt werden, und dem zu untersuchenden Bauelement. Als Probenhalter für die Messung dient der *Chuck* des Waferprobers, auf dem die Wafer mit Hilfe eines Vakuums gegen Verrutschen gesichert werden. Der *Chuck* lässt sich dabei für die Positionierung manuell oder halbautomatisch verfahren. Die Kontaktierung der Bauelemente erfolgt von oben über Wolfram-Nadeln, die direkt mit den SMUs verbunden sind, und von unten über den *Chuck* selbst. Der *Chuck* wird in dieser Arbeit üblicherweise als (*Back*)*Gate*-Kontakt genutzt, während die *Source/Drain*-Kontakte des Bauelements über die Nadeln kontaktiert werden.

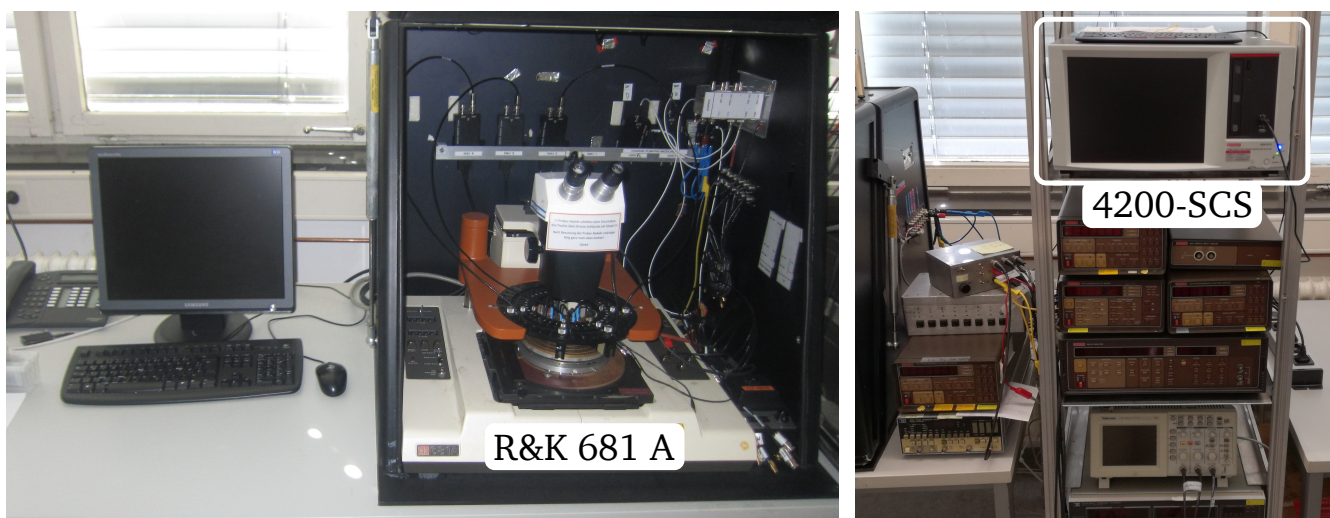


Abb. 18: Zur elektrischen Charakterisierung verwendeter Waferprober 681 A des Herstellers *Rucker and Kolls* mit angeschlossenem Messsystem 4200-SCS des Herstellers *Keithley* im Messlabor des IHTN.

5 Verwendete Prozesstechnologien

5.1 Reinigungsverfahren

Die Reinigungsverfahren für Silizium haben das Ziel, alle potentiellen Verunreinigungen von der Waferoberfläche zu entfernen. Auf Basis der Angaben des Herstellers wird davon ausgegangen, dass keine Kontaminationen des Wafervolumens vorliegen. Die möglichen Kontaminationen der Oberfläche werden in die drei Gruppen „Partikel“, „organisch“ und „metallisch“ eingeteilt. Nach KERN umfasst eine vollständige Reinigung die folgenden Schritte [49, Teil II, Kapitel 2]:

1. Entfernung organischer Kontaminationen
2. Entfernung des natürlichen Oxids
3. Entfernung von Partikeln und gleichzeitiges Neuwachsen eines Oxids
4. Entfernung metallischer Kontaminationen

Das natürliche Oxid, welches selbstständig auf einer Silizium-Oberfläche wächst, ist nicht *per se* als Verunreinigung zu betrachten. Es kann allerdings Kontaminationen *enthalten* und muss daher zunächst entfernt werden. Im Einzelnen wurden für die in dieser Arbeit verwendeten Wafer die folgenden Reinigungsschritte durchgeführt:

1. Partikel (grob):

Sofern die Wafer beschriftet oder auf andere Weise grobe Partikel erzeugt wurden, werden diese in einer Kaskade mit deionisiertem Wasser (DI-Wasser) abgespült, um keine makroskopischen Kontaminationen in die weiteren Bäder zu verschleppen.

2. Organik 1:

Organische Verunreinigungen stammen in der Regel vom Verpackungsmaterial und sollten zuerst entfernt werden, da sie die Oberfläche hydrophob machen und damit die Wirksamkeit anderer Reinigungslösungen reduzieren. Es wird dazu eine 1:10-Mischung aus Salpeter- und Schwefelsäure bei zunächst 110 °C und schließlich bei Raumtemperatur verwendet.

3. Organik 2:

In einem zweiten Schritt werden die noch verbliebenen organischen Verunreinigung mit einer 1:4-Mischung aus Wasserstoffperoxid und Schwefelsäure („*piranha clean*“) entfernt. Die Lösung reagiert beim Ansetzen stark exotherm und verbraucht sich durch Zersetzung des Wasserstoffperoxids innerhalb von 30 min.

4. natürliches Oxid 1:

Da die bisher verwendeten Lösungen einen stark oxidierenden Charakter haben, wurde das natürliche Oxid der Wafer noch nicht angegriffen. Es wird nun mit einer wässrigen 0,5 %igen Lösung der Flusssäure („*HF-Dip*“) entfernt. Dieser Arbeitsschritt ist kritisch, da die nun Wasserstoff-terminierte Waferoberfläche zwar sehr sauber, aber auch sehr anfällig gegen die Adsorption von Kohlenwasserstoffen und Partikeln ist.

5. Partikel (fein) & Passivierung:

Um eventuell adsorbierte Partikel zu entfernen und gleichzeitig die Oberfläche gegen weitere Partikel zu passivieren, wird eine auf 70 °C temperierte 20:4:1-Mischung von DI-Wasser, Wasserstoffperoxid und Ammoniumhydroxid („*standard clean 1*“ oder auch „*RCA clean 1*“) verwendet.

6. Metalle:

Zur Entfernung metallischer Kontaminationen wurde lange Zeit eine 5:1:1-Mischung aus DI-Wasser, Wasserstoffperoxid und Salzsäure („*standard clean 2*“ bzw. „*RCA clean 2*“) verwendet. Da aber auch verdünnte Mischungen metallische Verunreinigungen effektiv entfernen können [50] und gleichzeitig kostengünstiger sind [51, 52], wird hier eine stark verdünnte wässrige Lösung der Salzsäure verwendet, deren pH-Wert auf kleiner 2 eingestellt wird [49, Teil II, Kapitel 4].

7. natürliches Oxid 2:

Soll direkt im Anschluss eine Oxidation durchgeführt werden, wird das natürliche Oxid abermals mittels *HF-Dip* entfernt.



Abb. 19: Zur Wafer-Reinigung verwendete Nassbank mit DI-Wasserkaskade und Ätzbecken im Reinraum des IHTN.

Die unter Punkt 2 genannte Mischung aus Salpeter- und Schwefelsäure wird institutsintern als „Standard-Reinigung“ bezeichnet, ist aber mit keinem der beiden „*standard cleans*“ aus Punkt 5 oder 6 identisch. Mit Ausnahme dieser Mischung, die in den in Abb. 19 gezeigten Bädern der Nasschemie vorgehalten wird, werden alle anderen Reinigungslösungen jedes mal frisch angesetzt. Besonders zu beachten gilt, dass der *HF-Dip*, der im Fall einer direkt folgenden Oxidation zweifach benötigt wird, nach dem ersten Verwendung verworfen und neu angesetzt wird.

5.2 Thermische Oxidation

Hochwertige Gatedielektrika und Feldisolationschichten werden in der Halbleitertechnik durch thermische Oxidation von Silizium zu Siliziumdioxid hergestellt. Als Oxidant kommen eine Wasserdampf-Atmosphäre („Feuchtoxidation“, Gl. 31) oder reiner Sauerstoff („Trockenoxidation“, Gl. 32) zum Einsatz [53, Kapitel 3]. Die hinsichtlich elektrischem Durchbruch besseren Schichten liefert die Trockenoxidation, während die Feuchtoxidation höhere Aufwachsrate und damit dickere Schichten ermöglicht. Der Wasserdampf der Feuchtoxidation wird erzeugt, indem entweder Sauerstoff durch ein geheiztes Wasserbad geleitet wird („*bubbler*“) oder Wasserstoff und Sauerstoff in einer Knallgasflamme miteinander verbrannt werden.



Der Zusammenhang von Oxidationszeit und Dicke der gewachsenen Schicht in Abhängigkeit vom Oxidanten und der Temperatur wird durch das Modell von DEAL UND GROVE beschrieben, das die folgenden Prozesse zueinander in Beziehung setzt, wobei eine initial vorhandene dünne Oxidschicht vorausgesetzt wird [54]:

1. Transport des Oxidanten durch die Gasphase zur SiO_2 -Oberfläche
2. Diffusion des Oxidanten durch das bereits vorhandene Oxid zur SiO_2 -Si-Grenzfläche
3. Reaktion des Oxidanten an dieser Grenzfläche nach Gl. 31 oder Gl. 32 zu weiterem Oxid

Als Ergebnis liefert das Modell für kurze und lange Oxidationszeiten Gl. 33 bzw. Gl. 34 [55, Kapitel 5], über die die nötige Dauer für eine gewünschte Oxiddicke abgeschätzt werden kann. Dabei ist B die parabolische Wachstumskonstante, die die Diffusion des Oxidanten durch das Oxid beschreibt, und B/A die lineare Wachstumskonstante, die die Reaktionsrate an der SiO_2 -Si-Grenzfläche und den Gasfluss beschreibt. τ ist ein Korrekturterm, der der initial vorhandenen dünnen Oxidschicht Rechnung trägt.

$$d_{\text{Oxid}}^{\text{kurz}} = \frac{B}{A} \cdot (t + \tau) \quad (33)$$

$$d_{\text{Oxid}}^{\text{lang}} = \sqrt{B \cdot t} \quad (34)$$

Abb. 20 zeigt die in dieser Arbeit für die Oxidation verwendete Horizontal-Rohr-Ofenbank des Herstellers *Thermoco*. Die Rohre der Öfen sind in polykristallinem Silizium ausgeführt und erlauben die Oxidation von Substraten mit Durchmessern von maximal 3". Übliche Temperaturen liegen dabei im Bereich von 900 bis 1200 °C [53, Kapitel 3].

5.3 Photolithographie

Die Grundlage der Strukturübertragung bei der Herstellung integrierter Schaltungen bildet die Photolithographie. Den prinzipiellen Prozessablauf nach MACK zeigt Abb. 21 [56, Kapitel 1]: Zunächst wird das Schichtmaterial, das strukturiert werden soll, vollflächig auf dem gesamten Substrat abgeschieden („Beschichten“). Darauf wird ein photoempfindlicher Lack aufgebracht („Belacken“), über den durch „Belichten“ und „Entwickeln“ die gewünschte Struktur mit Hilfe einer Maske und einer geeigneten Ätztechnik in das Substrat übertragen wird. Nach dem „Übertragen“ wird der verbleibende Lack entfernt („Ablacken“).



Abb. 20: Zur thermischen Oxidation verwendete Horizontal-Rohr-Ofenbank des Herstellers *Thermoco* am IHTN; für Substrate mit Durchmessern von maximal 3".

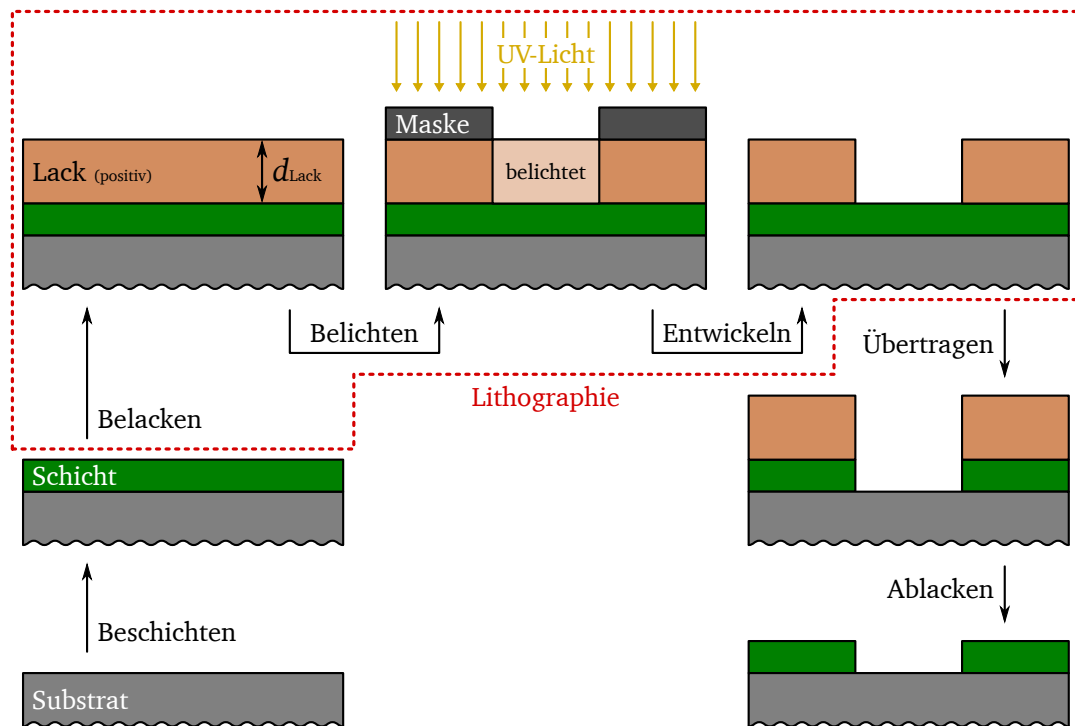


Abb. 21: Grundprinzip der Photolithographie zur Strukturübertragung nach [56, Kapitel 1] unter Verwendung eines Positiv-Lacks und der in dieser Arbeit genutzten Kontaktbelichtung.

Belacken

Vor dem eigentlichen Belacken werden die Substrate mindestens 1 h lang bei 100 bis 200 °C in einem Trockenschrank gelagert, um sämtliche Feuchtigkeit von der Oberfläche zu entfernen. Danach kann optional ein Haftvermittler aufgebracht werden, der die Haftungseigenschaften des Lacks verbessert. Haftvermittler können wie Lacke aufgeschleudert oder unter Vakuum über die Gasphase aufgedampft werden. Ein typischer Haftvermittler ist Hexamethyldisilazan (HMDS, $\text{C}_6\text{H}_{19}\text{NSi}_2$).

Lacke bestehen zumindest aus einem Lösungsmittel, über das die Viskosität eingestellt wird, einem Polymer, das als Matrix dient, und der eigentlichen photoempfindlich Komponente. Zum Aufbringen auf die gründlich gereinigten und getrockneten Wafer wird eine Lackschleuder („*spin coater*“, Abb. 22a) verwendet. Über die Drehzahl der Schleuder (üblicherweise zwischen 1000 und 6000 U/min) kann in Abhängigkeit von der Viskosität die Dicke der Lackschicht d_{Lack} eingestellt werden. Je viskoser ein Lack ist und/oder je langsamer er geschleudert wird, desto dicker ist die resultierende Lackschicht [56, Kapitel 1]. Um ein vorzeitiges Verdampfen des Lösungsmittels, und damit eine Zunahme der Viskosität, zu verhindern, sollte während des Belackens die Temperatur niedrig und konstant gehalten werden. Schließlich werden die belackten Wafer auf einer Heizplatte („*hot plate*“, Abb. 22b) oder einem Trockenschrank definiert getrocknet („*soft bake*“), um das noch vorhandene Lösungsmittel zu entfernen und den

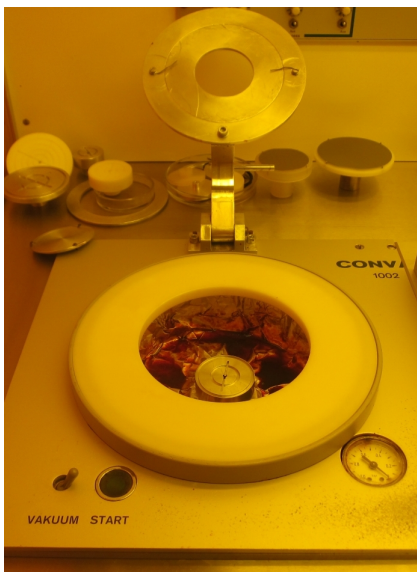
Lack für die Belichtung mittels eines Belichters („mask aligner“, Abb. 22c) mechanisch stabil zu machen [56, Kapitel 5].

Belichten

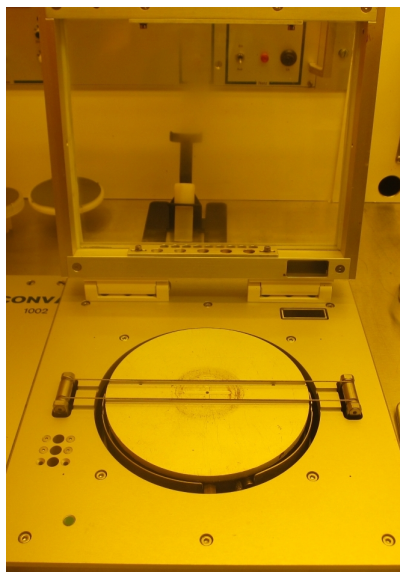
Beim Belichten werden zwei grundlegende Varianten unterschieden: *indirekte* Belichtung über eine Maske oder ein Retikel mit einer „Lichtquelle“ oder *direkte* Belichtung mit einem Elektronenstrahl. Bei der Belichtung über eine Maske wird Laser-Licht oder kurzwellige, breitbandige oder monochrome UV-Strahlung als „Lichtquelle“ verwendet. Für Auflösungen bis in den sub-Mikrometer-Bereich können zum Beispiel Quecksilberdampflampen eingesetzt werden. Es werden drei Varianten der indirekten Belichtung unterschieden [55, Kapitel 8]: die Proximitybelichtung, bei der ein Abstand von einigen Mikrometern zwischen Lack und Maske besteht; die in Abb. 21 skizzierte Kontaktbelichtung, bei der die Maske in direktem Kontakt mit dem Lack ist; und die Projektionsbelichtung, bei der zusätzlich zur Maske eine aufwendige Optik mit Linsen und Spiegeln benötigt wird.

Für die in dieser Arbeit verwendete Kontaktbelichtung ist die Auflösungsgrenze R durch Gl. 35 gegeben [55, Kapitel 8]. Der hier genutzte Belichter MA 56M des Herstellers Karl Süss nutzt als Lichtquelle eine Quecksilberdampflampe. Für die I-Linie des Spektrums dieser Lampe mit einer Wellenlänge λ von 365 nm und einer Lackdicke d_{Lack} von hier 1,4 μm ergibt sich eine Auflösungsgrenze von rund 700 nm.

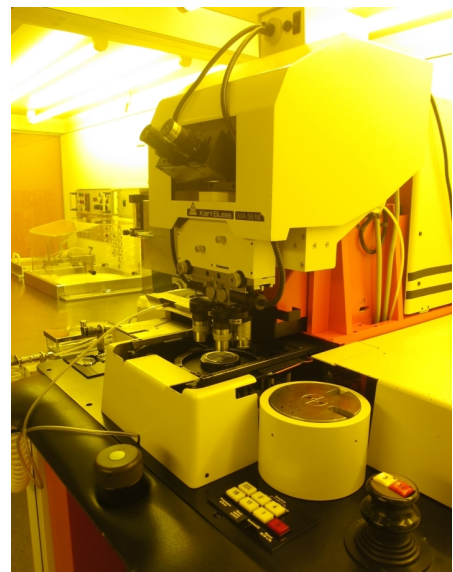
$$R_{\text{Kontakt}} = \sqrt{d_{\text{Lack}} \cdot \lambda} \quad (35)$$



(a) Lackschleuder („spin coater“) 1002 des Herstellers Convac



(b) Heizplatte („hot plate“)



(c) Belichter („mask aligner“) MA 56M des Herstellers Karl Süss

Abb. 22: Für die Photolithographie benötigte Geräte im Gelbraum des IHTN.

Entwickeln

Wurde beim Belacken ein „Positiv-Lack“ verwendet, so wurde durch die Belichtung eine Reaktion ausgelöst, die die Polymerketten des Matrixmaterials in den belichteten Bereichen teilweise auflöst und damit löslich gegen den Entwickler macht. Bei Verwendung eines „Negativ-Lacks“ wird dagegen eine Reaktion ausgelöst, die eine weitere Polymerisation der Matrix zur Folge hat und die belichteten Bereiche gegen den Entwickler stabilisiert, während die belichteten Bereiche löslich bleiben. Wird also bei gleicher Maske anstelle eines Positiv-Lacks ein Negativ-Lack verwendet, ergeben sich invertierte Strukturen [55, Kapitel 8] [56, Kapitel 7].

Nach dem Entwickeln kann, falls die Strukturübertragung durch Ätzen erfolgen soll, optional eine Härtung des Lacks durchgeführt werden („hard bake“). Diese Härtung findet wie die Lacktrocknung in einem Trockenschrank oder auf einer Heizplatte statt.

5.4 Strukturübertragung durch nasschemisches Ätzen

Nachdem die gewünschte Struktur durch Lithographie im Lack abgebildet wurde, wird sie durch eine nasschemische Ätzung in die darunter liegende Schicht übertragen. Nach der Strukturübertragung wird der noch auf den Substraten vorhandene Lack entfernt („Ablacken“). Dies kann, wie der eigentliche Ätzprozess selbst, entweder nasschemisch durch ein geeignetes Lösungsmittel oder trockenchemisch durch ein Sauerstoffplasma („Lackveraschen“) geschehen [53, Kapitel 2]. Beim nasschemischen Ätzen geschieht das Entfernen der Schicht dabei durch Eintauchen der Substrate in ein geeignetes, flüssiges Ätzmedium, das in einem Bad entweder vorgehalten oder jeweils frisch angesetzt wird. Durch das Medium werden die Bindungen zwischen den Atomen oder Molekülen der Schicht aufgebrochen, welche dann als Ionen oder Komplexe in gelöster Form vorliegen und über Diffusion und Konvektion von der Substrat-Oberfläche abtransportiert werden [57, Kapitel 15].

Nasschemische Ätzungen laufen in drei Schritten ab: Zunächst werden die Substrate in das Ätzmedium eingetaucht bis der gewünschte Ätzabtrag erfolgt ist. Danach werden die Substrate in einer Kaskade mit DI-Wasser gespült bis sämtliches Ätzmedium und gelöstes Schichtmaterial entfernt wurde. Zuletzt werden die Substrate getrocknet. Dies geschieht für komplette Hor-den automatisiert in einer Wafer-Schleuder oder für einzelne Wafer manuell mittels Stickstoff-Pistole. Die Dauer t des eigentlichen Ätzprozesses wird nach Gl. 36 durch die temperatur- und konzentrationsabhängige Ätzrate R des Mediums und den gewünschten Abtrag d bestimmt. Der zusätzliche Summand t_0 trägt Benetzungseffekten beim Eintauchen Rechnung.

$$t = R \cdot d + t_0 \tag{36}$$

5.5 Strukturübertragung mittels *Lift off*-Technik

Die sogenannte *Lift off*-Technik wird in dieser Arbeit zur Strukturierung des Katalysators und der *Source/Drain*-Kontakte verwendet und stellt eine Alternative zur Strukturübertragung durch Ätzung dar. Sie eignet sich besonders für Schichtmaterialien, die aufgrund fehlender Ätzmedien oder unzureichender Selektivität nicht geätzt werden können.

Verglichen mit der Strukturierung durch Ätzung werden bei der Strukturierung mittels *Lift off*-Technik die Schichtabscheidung und die Lithographie vertauscht: Bei der Strukturierung durch Ätzung wird erst die Schicht aufgebracht, dann die Lithographie durchgeführt und schließlich die Struktur durch Ätzen übertragen; bei der Strukturierung mittels *Lift off*-Technik dagegen, wird die Schicht auf den bereits belichteten und entwickelten Lack aufgebracht und die eigentliche Übertragung erfolgt dann durch Entfernen des Lacks und damit Abheben der auf dem Lack liegenden Schichtbereiche [57, Kapitel 15].

Damit ein *Lift off* möglich ist, darf der Lack keine senkrechten Flanken aufweisen, sondern es muss ein „Überhang“ oder „Unterschnitt“ erzeugt werden wie er in Abb. 23d skizziert ist. Die Erzeugung eines solchen Unterschnitts („*undercut*“) kann durch die Verwendung eines Doppellack-Systems mit einem zusätzlichen Copolymer oder durch die Verwendung eines Umkehrlacks („*image reversal resist*“) erreicht werden. In dieser Arbeit wurde ein Umkehrlack verwendet, nachdem sich das verfügbare Doppellack-System als prozesstechnisch schlecht beherrschbar herausgestellt hatte.

Der hier verwendete Umkehrlack hat folgende Eigenschaften: Er verhält sich bei Belichtung zunächst wie ein Positiv-Lack, d. h. belichtete Bereiche werden gegen den Entwickler löslich; lösliche Bereiche lassen sich thermisch in gegen den Entwickler stabile (inerte) Bereiche umkehren; weitere, zuvor nicht belichtete Bereiche reagieren nach wie vor wie ein Positiv-Lack, sodass sie durch einen zweiten Belichtungsschritt löslich gegen den Entwickler gemacht werden können. Im Einzelnen sind die folgenden in Abb. 23 skizzierten Arbeitsschritte für eine Strukturierung mittels *Lift off*-Technik nötig [58]:

a) Unterbelichtung:

Das belackte Substrat wird unter Verwendung einer Maske belichtet, die die gewünschten Strukturen als lichtundurchlässige Bereiche enthält („Stegmaske“). Die Belichtungs-dosis wird so gewählt, dass eine Unterbelichtung erzielt wird. Die Unterbelichtung hat zur Folge, dass nur im oberen Bereich des Lacks über die volle Breite der belichteten Bereiche eine Photoreaktion abläuft, während die Reaktion mit zunehmender Tiefe immer mehr auf die mittleren Bereiche beschränkt bleibt. Der effektiv belichtete Bereich ist an der Grenzfläche Lack-Substrat also schmaler als an der Oberkante, wodurch Flanken an den Rändern der Strukturen entstehen.

b) Umkehren:

Die Umkehrung von belichteten, löslichen Bereichen in inerte Bereiche erfolgt durch einen Umkehrbackschrift auf einer Heizplatte („*reversal bake*“). Dieser Schritt ist kritisch, da eine zu lange Dauer und/oder eine zu hohe Temperatur zu einer Quervernetzung des Lacks führt. Besonders bei ungenauer Kontrolle der Temperatur besteht die Gefahr eines unbeabsichtigten *hard bakes* des Lacks.

c) Flutbelichtung:

Die zweite Belichtung erfolgt ohne Maske mit hoher Dosis („*flood exposure*“), sodass sichergestellt ist, dass sämtliche zuvor nicht belichteten Bereiche nun belichtet und löslich gegen den Entwickler werden.

d) Entwickeln:

Durch die Entwicklung werden die nach unten schmaler werdenden Lackstege freigestellt. An dieser Stelle kann die Neigung der Lackflanken mittels Lichtmikroskop oder — aufwendiger und zerstörend, dafür auch präziser — mittels Rasterelektronenmikroskopie (SEM) kontrolliert werden.

e) Schichtabscheidung:

Auf den entwickelten Lack wird nun die Schicht abgeschieden. Das Abscheideverfahren sollte dabei möglichst gerichtet sein, damit keine vollständige Kantenbelegung erreicht wird und die unteren Bereiche der Lackflanke frei von Material bleiben. In diesem Prozessschritt muss darauf geachtet werden, dass die thermische Belastung des Lacks gering bleibt. Zu hohe Temperaturen können zu einem Verfließen der Unterschnitte oder zu einem Einbrennen des Lacks führen. Beides kann einen erfolgreichen *Lift off* erschweren oder gar vollständig verhindern. Aufgrund des begrenzten thermischen Budgets und der geforderten gerichteten Abscheidung kommen hier hauptsächlich PVD-Verfahren in Frage.

f) *Lift off*:

Die Abscheidung hat Schichtbereiche direkt auf dem Substrat und Bereiche auf dem Lack erzeugt, die wegen des Unterschnitts des Lacks von einer Lücke getrennt werden. In diese Lücke „kriecht“ in diesem Schritt der *Remover* und hebt durch Auf- und/oder Ablösen des Lacks die auf ihm liegenden Schichtbereiche ab. Zurück bleiben nur Schichtbereiche, die direkt auf dem Substrat abgeschieden wurden.

Zusammenfassend müssen also folgende allgemeine Voraussetzungen für einen erfolgreichen *Lift off* erfüllt werden:

- Die zu strukturierende Schicht ist dünn im Vergleich zum Lack.
- Der Unterschnitt ist groß im Vergleich zur Dicke der zu strukturierenden Schicht.
- Das Verfahren zur Schichtabscheidung ist möglichst gerichtet.
- Die thermische Belastung gering genug, dass kein Verfließen des Lacks stattfindet.

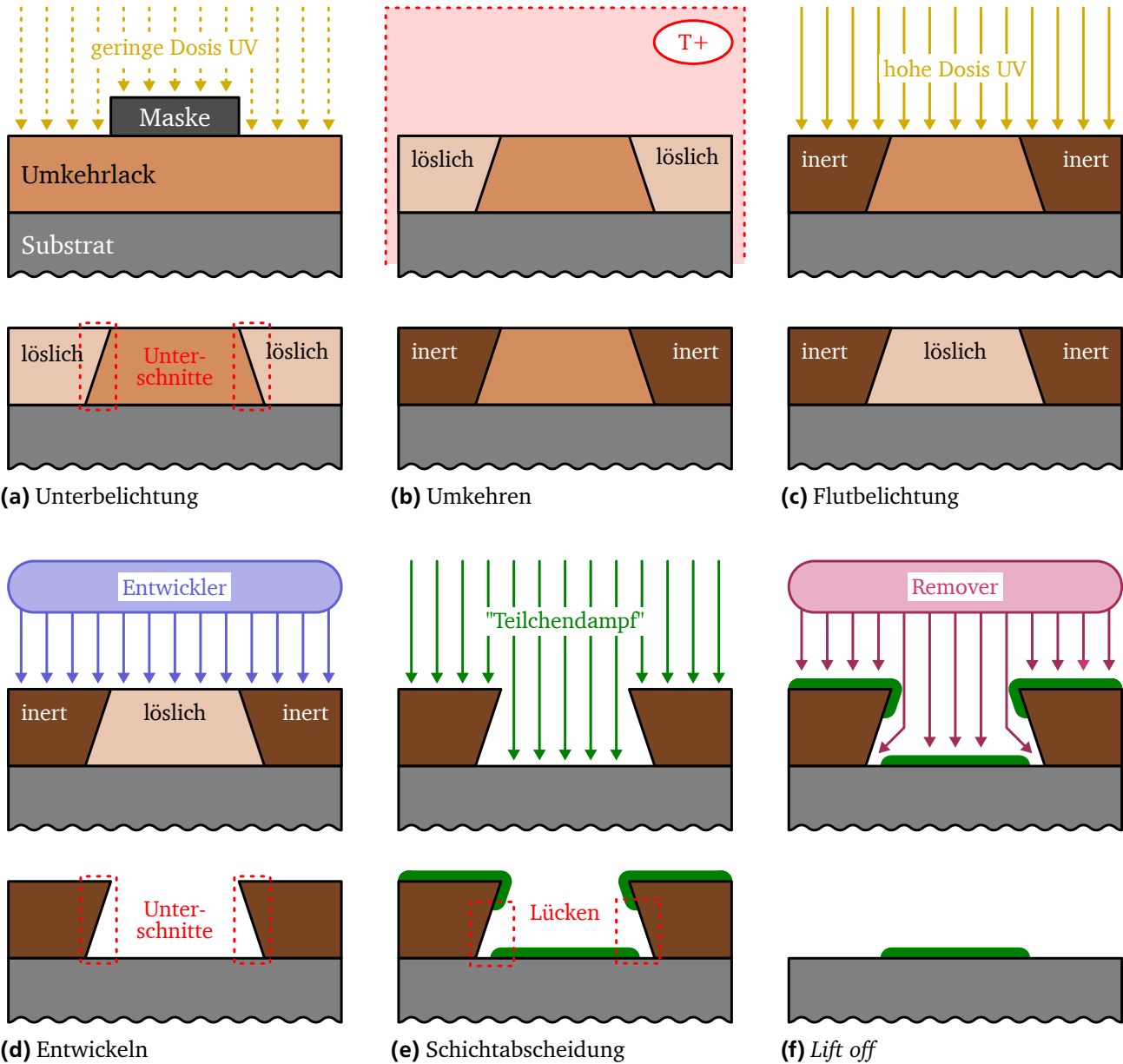
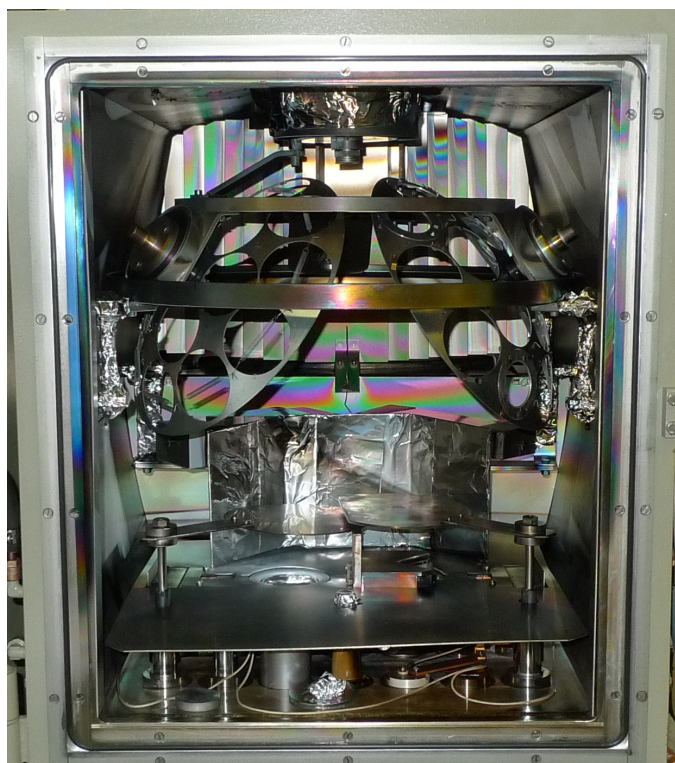


Abb. 23: Prinzip der *Lift off*-Technik; (a) bis (d) nach [58]: Durch Unterbelichtung, thermische Umkehrung, Flutbelichtung und Entwickeln werden Lackstege mit Unterschnitten an den Flanken erzeugt. Die gerichtete Schichtabscheidung kann die untersten Bereiche der Flanke nicht erreichen, wodurch Lücken in der Schicht entstehen, in die der *Remover* eindringen und das auf dem Lack liegende Material abheben kann.

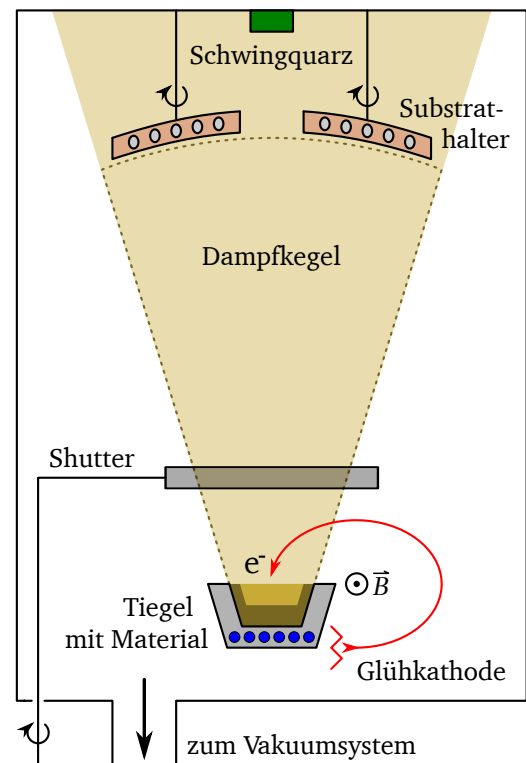
5.6 Elektronenstrahlverdampfung

Die Elektronenstrahlverdampfung stellt eine der drei Möglichkeiten zur thermischen Verdampfung dar, deren Übergruppe wiederum die physikalische Gasphasenabscheidung (PVD) ist. Verglichen mit den zwei anderen Varianten der thermischen Verdampfung, der Wendel- und Schiffchenverdampfung, werden mit der Elektronenstrahlverdampfung die Schichten mit der höchsten Reinheit erzielt. Das Verfahren basiert auf einer thermisch ausgelösten Verdampfung oder Sublimation des Aufdampfguts unter Hochvakuumbedingungen durch Beschuss mit einem Elektronenstrahl [57, Kapitel 10].

Abb. 24 zeigt den Rezipienten samt einer schematischen Darstellung der in dieser Arbeit verwendeten Anlage *BPU 100* des Herstellers *Balzers*. Der für die Verdampfung benötigte Elektronenstrahl wird erzeugt, indem von einer Glühkathode emittierte Elektronen durch eine Hochspannung von rund 10 kV beschleunigt werden [59, Kapitel 3]. Der erzeugte Elektronenstrahl wird durch ein Magnetfeld \vec{B} auf den Tiegel gelenkt, in dem sich das Aufdampfgut befindet. Eine Wasserkühlung des Tiegels stellt sicher, dass das Aufdampfgut niemals vollständig aufschmilzt, wodurch ein innerer „Eigenmaterial“-Tiegel entsteht, der die hohe Reinheit der Schichten begründet. Die Hochvakuumbedingungen von rund 10^{-6} mbar stellen sicher, dass



(a) geöffneter Rezipient



(b) Schema des Rezipienten

Abb. 24: Anlage *BPU 100* des Herstellers *Balzers* zur Elektronenstrahlverdampfung am IHTN, die zum Aufdampfen des Katalysators und zur Metallisierung verwendet wurde.

die Dampfpartikel eine genügend hohe freie Weglänge haben und sich damit ein gerichteter, ungestörter Dampfkegel bilden kann. In diesem Dampfkegel werden die zu beschichtenden Substrate positioniert, wo die Dampfpartikel zum Schichtmaterial kondensieren oder resublimieren. Um eine gleichmäßige Beschichtung aller Substrate zu gewährleisten, müssen alle Substrate im gleichen Abstand zum Tiegel positioniert werden. Diese Forderung wird durch die Verwendung eines sogenannten „Planetengeriebes“ als Substrathalter erfüllt, das durch mehrere in sich und gegeneinander rotierende Kugelsegmente gebildet wird [57, Kapitel 10].

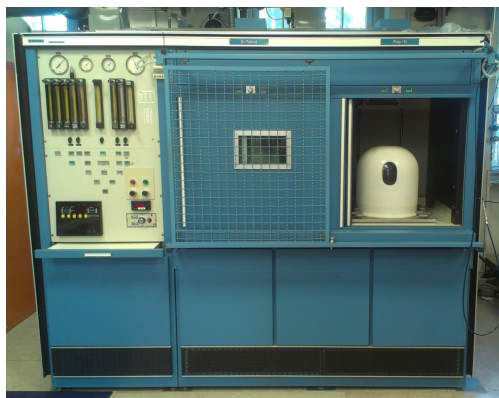
Die Dicke der abgeschiedenen Schicht wird kontinuierlich mit einer Quarzkristall-Mikrowaage (QCM, *quartz crystal microbalance*) bestimmt, die die masseabhängige Verstimmung eines schwingenden Quarzes ausnutzt. Nach SAUERBREY ist die Verstimmung des Schwingquarzes Δf über die Zunahme der Massenbelegung $\Delta m/A$ und eine Konstante des Schwingquarzes C_Q gegeben [60]. Wird die Zunahme der Massenbelegung allein durch die Masse m_S der abgeschiedenen Schicht bestimmt, so kann über die Verstimmung des Quarzes und die Dichte ρ_S der abgeschiedenen Schicht die Schichtdicke d_S nach Gl. 37 berechnet werden. Üblicherweise hat der Quarz aus konstruktiven Gründen einen größeren Abstand zum Tiegel als die Substrate, so dass ein Anlagenfaktor („*tooling factor*“) nötig wird, um die vermeintlich niedrigere Schichtdicke zu korrigieren.

$$\Delta f = -C_Q \cdot \frac{m_S}{A} = -C_Q \cdot \rho_S \cdot d_S \quad \text{mit} \quad m_S = \rho_S \cdot d_S \cdot A \quad (37)$$

5.7 Chemische Gasphasenabscheidung

Die chemische Gasphasenabscheidung (CVD) ist ein Verfahren zur Beschichtung von Substraten. Die Schichten können dabei reine Metalle, einfache Verbindungen wie Siliziumdioxid oder Siliziumnitrid, aber auch komplexere Strukturen wie der Hochtemperatursupraleiter $\text{YBa}_2\text{Cu}_3\text{O}_7$ sein. Nach CHOY „umfasst die CVD die Zersetzung und/oder chemische Reaktion von gasförmigen Reaktanden [...], die die Bildung eines stabilen, festen Produkts zur Folge hat“ [61]. Nach dieser allgemeinen Definition umfasst die CVD also ausdrücklich auch Reaktionen in der Gasphase, die zum Beispiel zur Herstellung von Pulvern genutzt werden können. In der Halbleitertechnik sind solche Reaktionen allerdings generell unerwünscht, da die Schichteigenschaften und die Homogenität der Abscheidung unzureichend sind.

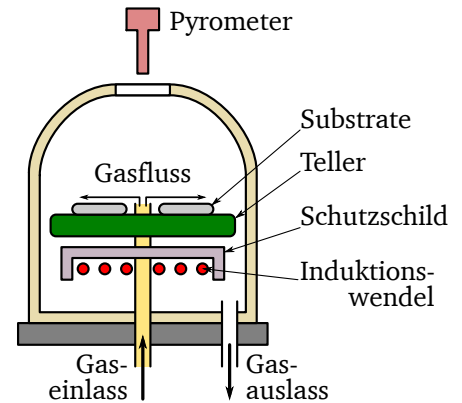
Die Aktivierungsenergie für die chemische Reaktion bzw. die Zersetzung wird meist durch Prozesstemperaturen von bis zu 1300 °C aufgebracht. Sind nur niedrigere Temperaturen möglich, kommen auch Plasma-unterstützte Prozesse (PE-CVD, *plasma-enhanced chemical vapor deposition*) oder Prozesse bei verringertem Druck (LP-CVD, *low-pressure chemical vapor deposition*) zum



(a) Gesamtansicht



(b) Reaktionskammer



(c) Schema der Reaktionskammer

Abb. 25: CVD-Anlage AVM 1200 des Herstellers *Applied Materials* am IHTN, die zur Cluster-Bildung und zum Wachsen der CNTs verwendet wurde.

Einsatz. CVD-Anlagen bestehen aus drei Komponenten: einer Gaszuführung, der Reaktionskammer und einem Abgassystem. Anlagen für den industriellen Einsatz können dabei im Grunde den gleichen Aufbau haben wie die Rohr-Öfen, die für die Oxidation zum Einsatz kommen. Kompaktere „pancake“-Reaktoren finden wegen ihres geringen Durchsatzes heute nur noch in der Forschung Verwendung.

Die in dieser Arbeit genutzte Anlage AVM 1200 des Herstellers *Applied Materials* ist in Abb. 25a in einer Gesamtansicht gezeigt. Die geöffnete Reaktionskammer ist in Abb. 25b abgebildet und in Abb. 25c skizziert. Die Kammer besteht aus einer Quarz-Glocke und einem rotierenden mit Siliziumkarbid beschichteten Graphit-Teller. Durch den Teller, der als Substrathalter dient, ist mittig der Gaseinlass geführt. Der Teller wird induktiv durch eine Hochfrequenz geheizt, die durch eine unter dem Teller befindliche Spule eingekoppelt wird. Um Teller und Spule räumlich von einander zu trennen, ist zwischen ihnen ein „Schutzschild“ aus Quarzglas montiert.

Die symbolischen Reaktionsgleichungen für die Abscheidung einer metallischen Schicht M durch Zersetzung und für die Abscheidung einer Verbindung AB durch chemische Reaktion sind in Gl. 38 und Gl. 39 gegeben. In Gl. 40 ist nach JAEGER als Beispiel für eine konkrete Reaktionsgleichung die Abscheidung von Siliziumnitrid mit Hilfe von Silan (SiH_4) und Ammoniak (NH_3) gegeben [53, Kapitel 6].



Der prinzipielle Ablauf der Reaktionen für die Abscheidung einer Schicht ist nach PIERSON in Abb. 26 skizziert [62, Kapitel 2]. Im Einzelnen laufen folgende Schritte ab:

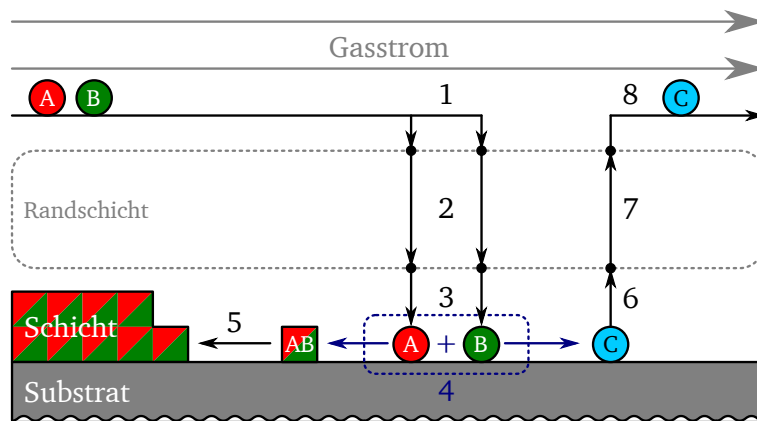


Abb. 26: Prinzipieller Ablauf eines CVD-Prozesses am Beispiel einer Reaktion mit zwei Edukten zu einer binären Verbindung nach [62, Kapitel 2]: Nachdem die Edukte A und B die Oberfläche erreicht haben, reagieren sie zum Schichtmaterial AB, das an der Oberfläche zu einem „Einbauplatz“ diffundiert, während die Co-Produkte C über die Gasphase entfernt werden.

1. Antransport der Edukte über den Gasstrom
2. Diffusion der Edukte durch die Randschicht zur Oberfläche
3. Adsorption der Edukte an die Oberfläche
4. Reaktion der Edukte zum Produkt
5. Oberflächendiffusion des Produkts zu einem „Einbauplatz“
6. Desorption der Co-Produkte von der Oberfläche
7. Diffusion der Co-Produkte durch die Randschicht zum Gasstrom
8. Abtransport der Co-Produkte durch den Gasstrom

5.8 Atomlagenabscheidung

Die Atomlagenabscheidung (ALD) ist ein Verfahren zur Schichtabscheidung, das in den 1970er-Jahren von SUNTOLA UND ANTSON ursprünglich unter der Bezeichnung „atomic layer epitaxy (ALE)“ entwickelt wurde [63]. Es handelt es sich hierbei um einen modifizierten CVD-Prozess, weswegen an einigen Stellen auch die Bezeichnung ALCVD (*atomic layer chemical vapour deposition*) Verwendung findet.

Ebenso wie bei der CVD werden bei der ALD zwei oder auch mehr verschiedene Edukte („Precursoren“) in einem Reaktor zur Reaktion gebracht. Die konkrete Anzahl der nötigen Precursoren ist meist durch die Anzahl der Elemente im gewünschten Schichtmaterial gegeben: für eine binäre Verbindung werden zwei Precursoren benötigt, für eine ternäre drei und so weiter. Die Eigenheit der ALD besteht darin, dass sich im Gegensatz zur CVD immer nur ein Precursor im Reaktor befindet. Statt sämtliche Precursoren gleichzeitig und kontinuierlich in den Reaktor zu leiten, wird bei der ALD also mit einem Zyklus von sequenziellen Precursor-Pulsen gearbeitet,

wie er auch in Abb. 27 skizziert ist. Zwischen den einzelnen Precursor-Pulsen wird der Reaktor mit einem Inertgas gespült („*purge*“), um eine Trennung der verschiedenen Precursoren sicherzustellen. Pro eingesetztem Precursor sind für einen vollständigen Zyklus also jeweils ein Arbeits- und ein Spülschritt nötig. Die entscheidende Folge dieser sequenziellen Arbeitsweise ist, dass die Reaktion der Precursoren zum Schichtmaterial in zwei separate „Halbreaktionen“ aufgeteilt wird. Durch diese Aufteilung wird die Reaktion, die bei der CVD theoretisch auch in der Gasphase stattfinden kann, zwingend auf die Oberfläche des Substrats verlagert.

Der Mechanismus der Schichtabscheidung mittels ALD kann wie folgt beschrieben werden: Das Substrat wird dem ersten Precursor ausgesetzt, wobei im Idealfall eine einzige Monolage davon auf der Substrat-Oberfläche chemisorbiert. Anschließend werden der Überschuss des Precursors und die Reaktionsprodukte der Chemisorption unter Einsatz des inerten Spülgases aus dem Reaktor entfernt. Dann wird das Substrat dem zweiten Precursor ausgesetzt, der mit der bereits chemisorbierten Monolage des ersten Precursors reagiert. Der Überschuss des zweiten Precursors und die Reaktionsprodukte werden wiederum durch Spülen entfernt. Damit ist ein Abscheidezyklus vollständig abgeschlossen und der nächste kann begonnen werden. Die an den Prozess gestellte Anforderung, dass dabei immer nur eine einzige Monolage des jeweiligen Precursors chemisorbiert bzw. reagiert wird als „Selbstlimitation“ bezeichnet und ist die wichtigste Eigenschaft, die ein Precursor für einen ALD-Prozess erfüllen muss. Sind alle Prozessparameter optimal gewählt, so ist das pro Zyklus erzielte Wachstum der Schicht (*GPC*, *growth per cycle*) konstant. Der *GPC*-Wert stellt eine wichtige Kenngröße des jeweiligen ALD-Prozesses dar und liegt bei durchschnittlichen Prozessen im unteren einstelligen Angström-Bereich pro Zyklus.

Bei der in dieser Arbeit genutzten ALD-Anlage handelt es sich um die *Savannah S100* des Herstellers *UltraTech / Cambridge NanoTech*. Die in Abb. 28a abgebildete Anlage steht im Reinraum der TECHNISCHEN UNIVERSITÄT BERGAKADEMIE FREIBERG; die Arbeiten daran wurden somit extern durchgeführt. Die Anlage besteht im Wesentlichen aus drei Komponenten, die auch in Abb. 28b skizziert sind: den sogenannten „Linien“, die die Precursoren enthalten, einem Re-

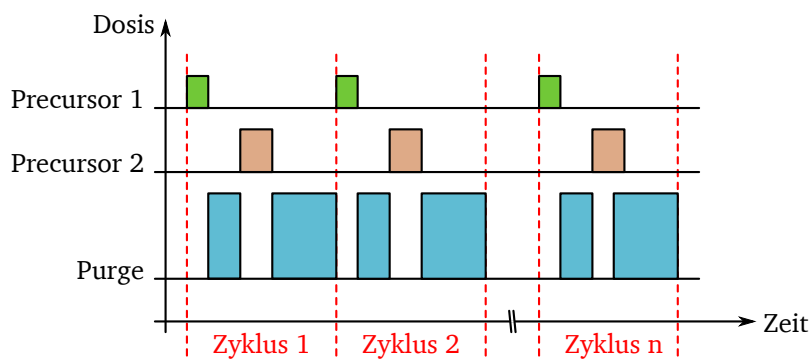
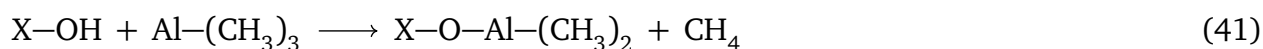


Abb. 27: ALD-Zyklus eines Prozesses mit zwei Precursoren nach [64]: Um eine Trennung der Precursoren im Reaktor sicherzustellen, werden zwischen die eigentlichen Arbeitsschritte jeweils Spülschritte eingefügt, sodass ein vollständiger Zyklus aus den beiden Precursor-Pulsen und zwei *purge*-Pulsen besteht.

aktor mit heizbarem Substrathalter und einem Pumpensystem. Pro verwendetem Precursor ist eine Linie nötig, wobei sich die verschiedenen Linien je nach den Eigenschaften der Precursoren in ihrer Beschaffenheit unterscheiden können. Das Pumpensystem dient der Einstellung des Basisdrucks in der gesamten Anlage und zur Entfernung von überschüssigem Precursor und der Reaktionsnebenprodukte. Für alle Zuleitungen und den Reaktor selbst müssen Materialien verwendet werden, die durch die Precursoren nicht angegriffen werden und von denen eine schnelle Desorption aller auftretender Verbindungen stattfindet.

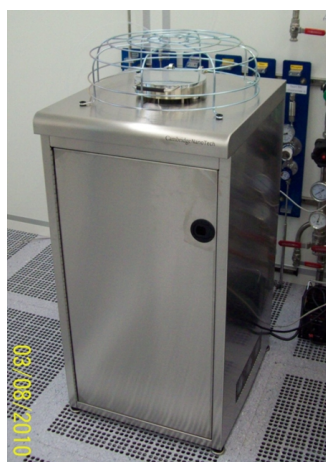
In dieser Arbeit wird die ALD zur Abscheidung von Aluminiumoxid genutzt, wobei Trimethylaluminium (TMA) und DI-Wasser als Precursoren zum Einsatz kommen. Der TMA/Wasser-Prozess stellt im Allgemeinen einen der am meisten untersuchten ALD-Prozesse und im Besonderen die einfachste und robusteste Möglichkeit für die Abscheidung von Aluminiumoxid dar [65, 66]. TMA dient hier als Aluminium-Quelle, während durch das Wasser der Sauerstoff zur Verfügung gestellt wird. Alternativ können auch Wasserstoffperoxid [67, 68] oder reiner Sauerstoff bzw. Aluminiumchlorid [69] genutzt werden.

Die Schichtbildung läuft während des TMA/Wasser-Prozesses gemäß der beiden Halbreaktion in Gl. 41 und Gl. 42 und wie in Abb. 29 skizziert ab, wobei für die Ausgangssituation $X=Si$ und nach dem ersten Zyklus $X=Al$ gilt:

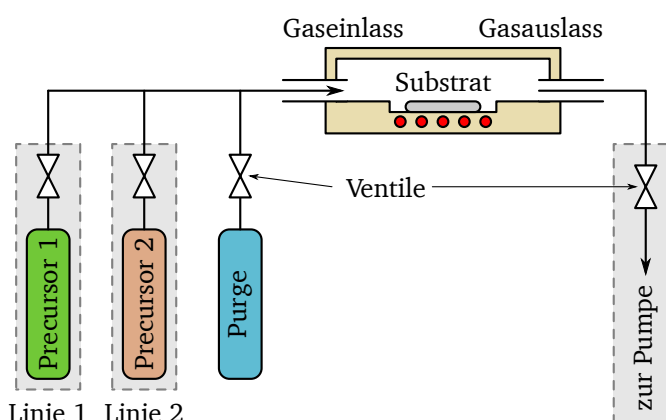


a) Ausgangssituation:

Zunächst wird davon ausgegangen, dass das Substrat mit OH-Gruppen terminiert ist. Mit Blick auf die Reinigung der Substrate bedeutet dies, dass auf einen *HF-Dip* verzichtet oder



(a) Gesamtansicht



(b) schematische Darstellung

Abb. 28: ALD-Anlage *Savannah S100* der Firma *UltraTech / Cambridge NanoTech* im Reinraum der TECHNISCHE UNIVERSITÄT BERGAKADEMIE FREIBERG, die für die Abscheidung von Aluminiumoxid genutzt wurde.

gewartet werden muss, bis die Oberfläche rehydriert ist. Andernfalls wird die durch den *HF-Dip* erzeugte Wasserstoff-Terminierung den hier gewählten Prozess empfindlich stören.

b) erste Halbreaktion:

Während des ersten TMA-Pulses reagieren eine oder zwei OH-Gruppen der Substratoberfläche unter Abspaltung einer oder zwei Methyl-Gruppen eines TMA-Moleküls, wobei Methan (CH_4) als Nebenprodukt frei wird. Diese Reaktion läuft so lang ab, bis sämtliche OH-Gruppen entfernt wurden und eine Methyl-Terminierung vorliegt, die gegen weiteres TMA inert ist. Überschüssiges TMA und das Nebenprodukt Methan werden dann in der Spülphase aus dem Reaktor entfernt.

c) zweite Halbreaktion:

Der Wasser-Puls entfernt nun die Methyl-Gruppen, wiederum unter Bildung von Methan als Nebenprodukt, wodurch wieder eine Terminierung der Oberfläche mit OH-Gruppen vorliegt, die gegen weiteres Wasser inert ist. Überschüssiges Wasser und das Methan werden dann in der Spülphase aus dem Reaktor entfernt, womit die Probe für den nächsten Zyklus mit dem nächsten TMA-Puls bereit ist.

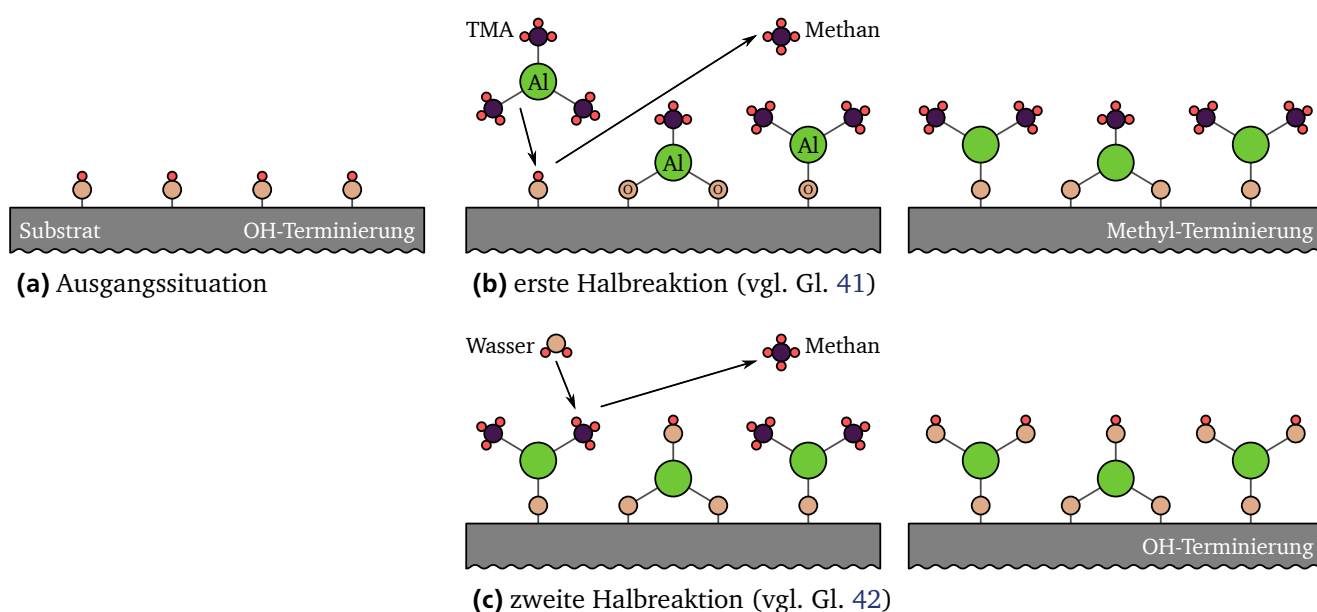


Abb. 29: Prinzip der ALD am Beispiel des hier genutzten TMA/Wasser-Prozesses zur Abscheidung von Aluminiumoxid [65]: Ausgehend von einem OH-terminierten Substrat **(a)** werden in zwei Halbreaktionen die beiden Elemente der Schicht sequenziell zur Verfügung gestellt. In der ersten Halbreaktion **(b)** chemisorbiert das TMA an den OH-Gruppen unter Bildung von Methan, wobei sich eine Methyl-Terminierung der Oberfläche einstellt. Durch die zweite Halbreaktion **(c)** werden die Methyl-Gruppen durch Wasser entfernt und wieder durch OH-Gruppen ersetzt, sodass auf der Oberfläche wieder die Ausgangssituation hergestellt ist. Beide Halbreaktionen sind selbst limitierend, da sowohl eine Methyl-terminierte als auch eine OH-terminierte Oberfläche gegen weiteres TMA bzw. Wasser inert sind.

5.9 Reaktives Sputtern

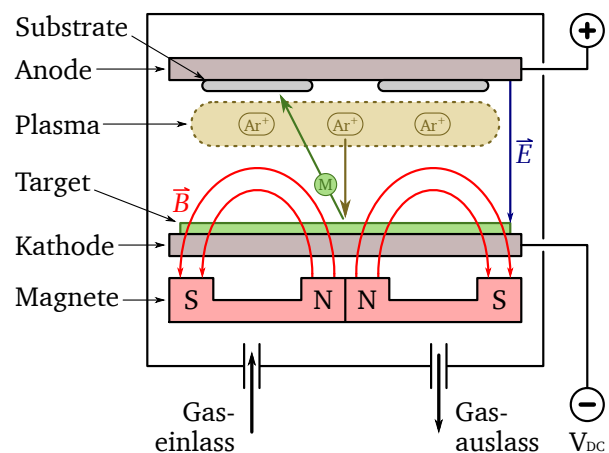
Das *Sputtern* (eigentlich „Kathodenzerstäuben“, auch „*Sputterdeposition*“) stellt neben der Elektronenstrahlverdampfung ein weiteres PVD-Verfahren dar. Dabei befindet sich das abzuscheidende Material als sogenanntes „*Target*“ auf der Kathode eines Plattenkondensators. Beim *Sputtern* wird unter Vakuumbedingungen zwischen den beiden Elektroden des Kondensators ein Argon-Plasma erzeugt und die im Plasma enthaltenen Argon-Ionen in Richtung der Kathode beschleunigt. Durch den Eintrag der kinetischen Energie der beschleunigten Ionen wird so das *Target* zerstäubt [59, Kapitel 3]. Der auf diese Weise erzeugte Teilchendampf schlägt sich schließlich auf den auf der Anode befindlichen Substraten nieder. Dadurch, dass beim *Sputtern* das Material nicht erst aufschmilzt und dann verdampft, sondern direkt in festem Zustand in die Gasphase übergeht, können auch nicht kongruent schmelzende Verbindungen gesputtert werden. Damit ist auch eine Abscheidung von Legierungen und chemischen Verbindungen möglich, die sich bei der Elektronenstrahlverdampfung üblicherweise entmischen bzw. zersetzen würden.

Von „reaktivem *Sputtern*“ wird gesprochen, wenn dem Sputtergas noch ein Reaktivgas beige-mischt wird. Das Reaktivgas soll dabei nicht die Zerstäubung des *Targets* unterstützen, sondern mit dem erzeugten Teilchendampf reagieren. So können auf einfache Weise beispielsweise Metalloxide abgeschieden werden, indem reine Metall-*Targets* und Sauerstoff als Reaktivgas genutzt werden.

In dieser Arbeit wird für die Abscheidung der alternativen Katalysatoren eine Anlage zum Gleichstrom-Magnetronsputtern („*DC magnetron sputtering*“) verwendet. Der Rezipient der Anlage ist in Abb. 30a gezeigt und das Prinzip der Schichtabscheidung in Abb. 30b schematisch dargestellt. Im Wesentlichen besteht die Anlage aus einem Plattenkondensator, einer Gaszu-



(a) geöffneter Rezipient



(b) Prinzip des DC-Magnetronsputterns nach [70]

Abb. 30: Gleichstrom-Magnetronsputter-Anlage am IHTN (mit *S-Gun 980-2060* des Herstellers *Varian* modifizierte *BPU 100* des Herstellers *Balzers*), die zum Abscheiden der alternativen Katalysatoren verwendet wurde.

führung für das Sputter- und das Reaktivgas und einem Pumpensystem für die Erzeugung eines Vakuums in der Größenordnung von 10^{-2} Torr. Über Anlegen einer Gleichspannung V_{DC} an den Plattenkondensator wird in ihm über das elektrische Feld \vec{E} das Plasma gezündet und die erzeugten Ionen beschleunigt. Zusätzlich befinden sich unter der Kathode mit dem *Target* Magnete, deren Magnetfeld \vec{B} die von der Kathode emittierten Elektronen auf Spiralbahnen lenkt, was den Ionisationsgrad des Sputtergases steigert; auf diese Weise wird die Sputter- und damit auch der Abscheiderate erhöht [57, Kapitel 10].

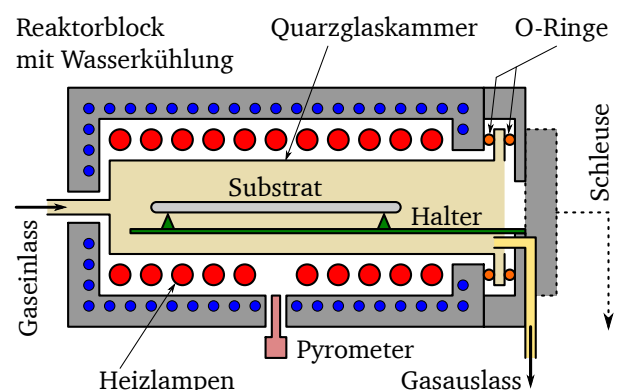
5.10 Kurzzeit-Prozessierung

Die Kurzzeit-Prozessierung (RTP) umfasst zwei Varianten: die schnelle thermische Oxidation (RTO) und die schnelle thermische Ausheilung (RTA). Die RTO ist ein Verfahren zum Aufwachsen extrem dünner Oxide auf Silizium ($d < 4\text{ nm}$ [71]), während die RTA genutzt wird um Kristalldefekte in Silizium nach Implantationsprozessen zu reparieren. Anlagentechnisch gesehen unterscheiden sich RTO- und RTA-Prozesse nur hinsichtlich der angeschlossenen Gase. Beide RTP-Varianten erfordern hohe Heizraten in der Größenordnung von bis zu 100 K/s [72] und Prozesszeiten von unter 1 min [71] bei einer gleichzeitig präzisen Kontrolle der Temperatur.

RTP-Anlagen bestehen im Wesentlichen aus einer Gaszuführung und einem Reaktorblock samt Heizer. Da die Wafer in RTP-Anlagen einzeln prozessiert werden, ist in der Regel noch ein Roboter zur Beschickung der Anlage vorhanden, um den Durchsatz zu erhöhen. Eine Frontalansicht der in dieser Arbeit für die Cluster-Bildung verwendeten Anlage *SHS 1000G* des Herstellers *AST* zeigt Abb. 31a. Der wassergekühlten Reaktorblock ist in Abb. 31b skizziert. In ihm befinden sich eine obere und eine untere Reihe von Halogenlampen, die als Heizer dienen, und dazwischen eine Kammer aus Quarzglas, in der sich während des Prozesses der Wafer befindet. In der Kammer wird der Wafer auf kleinen Pins abgelegt, um seinen thermischen Kontakt zur Umgebung



(a) Frontalansicht mit Ladestation



(b) Schema des Reaktorblocks nach [73]

Abb. 31: RTP-Anlage *SHS 1000G* des Herstellers *AST* am IHTN, die für die Cluster-Bildung verwendet wurde.

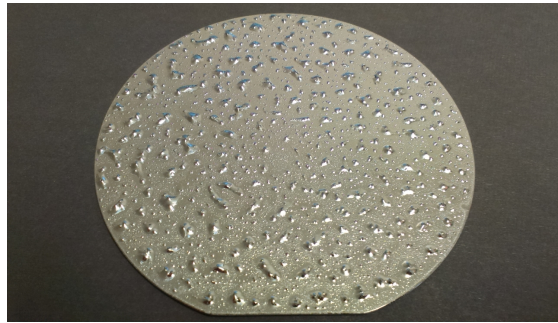


Abb. 32: Durch Ausfall der RTP-Temperatur-Regelung oberflächlich aufgeschmolzener 4''-Wafer.

so gering wie möglich zu halten. Um auch den Kontakt der Kammer nach außen hin zu minimieren, ist sie lediglich an der Frontseite mit einem Paar O-Ringen befestigt. Zur Messung der Temperatur wird ein Pyrometer verwendet, das von unten durch die Kammer hindurch auf die Rückseite des Wafer gerichtet ist.

Um die geforderten hohen Heizraten zu realisieren, sind die Heizlampen auf eine Leistung von über 1 kW pro Lampe ausgelegt [73]. Die damit unter Volllast über längere Zeit erreichbaren Temperaturen übersteigen ebenso die Belastbarkeit der Anlage wie auch den Schmelzpunkt von Silizium. (Das Ergebnis eines „Störfalls“ zeigt Abb. 32: Wie deutlich zu erkennen ist, ist die zuvor polierte, spiegelnde Oberfläche des Wafer nun rau und von Blasen überzogen.)

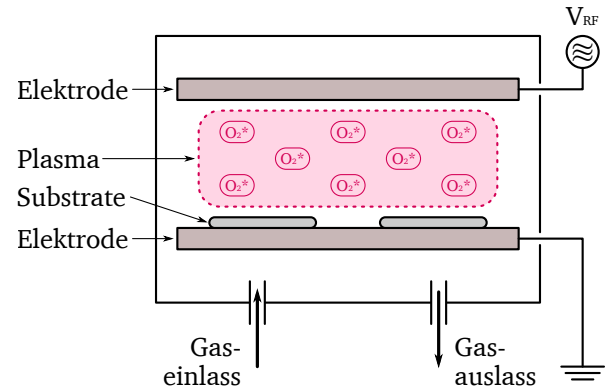
5.11 Sauerstoff-Plasmabehandlung

Bei der in dieser Arbeit genutzten Sauerstoff-Plasmabehandlung handelt es sich um einen Prozess, der am Institut ursprünglich zur Entfernung des Lacks nach der Strukturübertragung durch Ätzen oder nach Ionenimplantationen verwendet wurde. Dieses sogenannte „Lackveraschen“ (auch „*O₂-Dip*“) stellt einen trockenchemischen Plasma-Ätzprozess dar (CDE, *chemical dry etching*). CDE-Prozesse sind üblicherweise isotrop und haben eine hohe Selektivität [59, Kapitel 5].

Der in dieser Arbeit genutzte Trockenätzter ist in Abb. 33a gezeigt und das Prinzip eines CDE-Prozesses in Abb. 33b am Beispiel des Lackveraschens skizziert. Dabei wird bei einem Druck im Bereich von 0,1 bis 5 Torr in einem Kondensator mit symmetrischen Elektroden ein Sauerstoff-Plasma erzeugt, das freie Radikale („O₂^{*}“ in Abb. 33b) enthält [53, Kapitel 2]. Diejenige Elektrode, auf der sich auch die Substrate befinden, wird auf Erdpotential gehalten, während an die andere Elektrode die für die Zündung des Plasmas benötigte hochfrequenten Wechselspannung V_{RF} angelegt wird [57, Kapitel 16]. Die im Plasma enthaltenen hochreaktiven Sauerstoff-Radikale oxidieren den Lack, wodurch er in die Gasphase übergeht und schließlich über das Pumpensystem entfernt wird.



(a) Frontalansicht



(b) Prinzip des CDE nach [57, Kapitel 16]

Abb. 33: Für die Sauerstoff-Plasmabehandlung des Katalysators verwendeter Trockenätzter am IHTN (Eigenbau des Instituts mit Plasmaquelle des Herstellers *Hüttinger*).

5.12 Ozongenerator

Für die Nachoxidation der gesputterten Aluminiumoxidschichten wird in dieser Arbeit der in Abb. 34 abgebildete selbstgebaute Ozongenerator genutzt. Der Generator besteht neben einem Hochspannungsnetzteil aus einem Kondensator und einer Pumpe, die über Silikonschläuche miteinander verbunden sind. Im Betrieb wird, nachdem die Netzspannung mittels des verbauten Netzteils hochtransformiert wurde, im Kondensator Luftsauerstoff durch Elektrophorese in Ozon umgewandelt. Das erzeugte Ozon wird mit Hilfe eines speziellen Probenhalters über die Wafer und schließlich in die Abluft des Labors geleitet. Der Probenhalter kann in einen *Mask aligner* eingelegt werden, sodass eine UV-unterstützte Nachoxidation möglich ist.

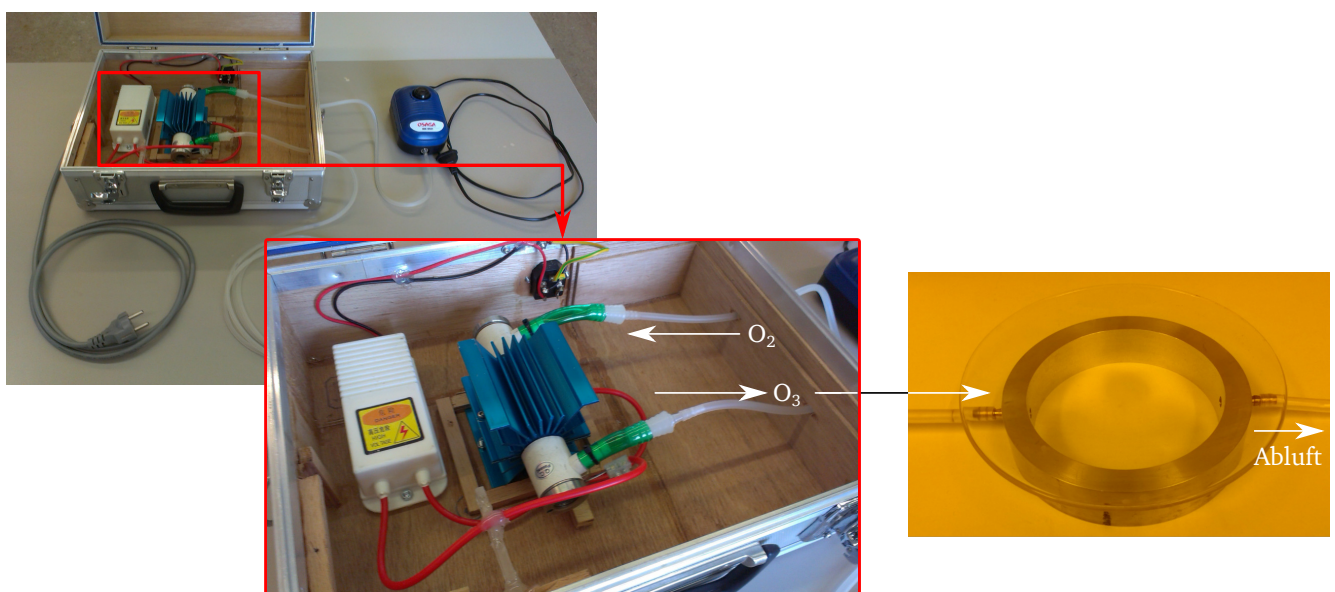


Abb. 34: Für die Nachoxidation der gesputterten Aluminiumoxidschichten genutzter Ozongenerator.

6 Prozess zur Herstellung von CNTFETs

Im Folgenden soll der technologische Prozess zur Herstellung von Kohlenstoffnanoröhren-Feldeffekt-Transistoren (CNTFETs) beschrieben werden, der in dieser Arbeit ausgehend von dem in der Vorgängerarbeit [6] beschriebenen Verfahren entwickelt wurde. Die Grundlage bildet nach wie vor das *in situ*-Wachstum von CNTs durch die katalytisch-chemische Gasphasenabscheidung (CCVD) mit Hilfe von nanoskaligen Nickel-Clustern und Methan. Die Cluster werden zuvor durch einen Temperschritt aus einem metallischen Doppelstapel von Aluminium und Nickel erzeugt. Während der Temperung erfahren die Metallschichten eine Transformation: die konforme Nickel-Schicht wird in diskrete Cluster und das metallische Aluminium in ein dielektrisches Aluminiumoxid umgewandelt, das die später gewachsenen CNTs elektrisch voneinander isoliert. Eine Neuerung gegenüber der Vorgängerarbeit stellen die Einführung des „strukturierten Katalysators“ und eines anderen Testchip-Layouts dar.

Wie in der Vorgängerarbeit wird hier bewusst auf die Herstellung individueller *Topgates* verzichtet. Statt also für jeden auf einem Substrat hergestellten Transistor ein eigenes *Gate* zur Verfügung zu stellen, wird in dieser Arbeit ein Ansatz verfolgt, der für alle Bauelemente den Silizium-Wafer als gemeinsames *Backgate* nutzt. Dieser Ansatz ermöglicht es, zu jedem Zeitpunkt des Prozesses topologische Messung durchführen zu können. Dies wird, ebenso wie die für heutige Verhältnisse sehr langen Kanäle der Transistoren im Mikrometerbereich und das dicke Siliziumdioxid-*Backgate*, die Leistungsfähigkeit der CNTFETs hinsichtlich des Stromes im *On*-Zustand und hinsichtlich des Schaltverhaltens einschränken. Diese Einschränkungen werden hier zu Gunsten der Möglichkeit in Kauf genommen, auch an vollständig prozessierten Transistoren noch topologische Messungen mittels AFM durchführen zu können.

Anmerkung zum Begriff des „Katalysators“

Der Begriff des „Katalysators“ wird in dieser Arbeit — wie in diesem Technologiefeld üblich — weiter gefasst gebraucht. Am zutreffendsten wäre es, die Nickel-Cluster als Katalysator zu bezeichnen, da von ihnen aus das Wachstum der CNTs beginnt. Aber auch die Cluster entsprechen nicht der eigentlichen Definition eines Katalysators, da an oder in ihnen Rückstände von Kohlenstoff bleiben und das Wachstum der CNTs im engeren Sinne keine chemische Reaktion darstellt. Aus diesem Grund soll unter „Katalysator“ hier schlicht sowohl der metallische Doppelstapel von Aluminium und Nickel *vor* der Temperung als auch das umgewandelte Aluminiumoxid samt der Nickel-Cluster *nach* der Temperung verstanden werden.

6.1 Testchip-Layout

Das für diese Arbeit entwickelte Testchip-Layout für 2"-Wafer ist in Abb. 35 gezeigt. Es enthält große, mit bloßem Auge sichtbare Strukturen — Pfeile, Flatjustage und Achsen (gelbe Markierungen) — für eine grobe Justage beim Belichten und eine allgemeine Orientierung bei der Handhabung der Wafer. Vom Pinzettenrand abgesehen, der weitestgehend frei gehalten wird, sind alle nicht von Strukturen belegten Bereiche mit einem Schachbrettmuster gefüllt, um große zusammenhängende Flächen zu vermeiden, die während der *Lift off*-Prozesse störend sein können. Die Justagemarken (violett) sind paarweise im linken und rechten Bereich des Wafers an der horizontalen Achse positioniert.

Den größten Raum im Layout nehmen die acht Felder mit Fingertransistoren (rot) ein, die für die eigentlichen Messungen verwendet wurden. Die vorgesehen Kanallängen und -weiten dieser Transistoren sind in Tab. 2 (Seite 59) aufgelistet. Neben den genannten sind noch weitere Strukturen in diesem Testchip-Layout vorhanden, die in dieser Arbeit aber schlussendlich nicht genutzt wurden. Eine Übersicht dieser Strukturen ist ab Seite 129 in den Anhang gestellt.

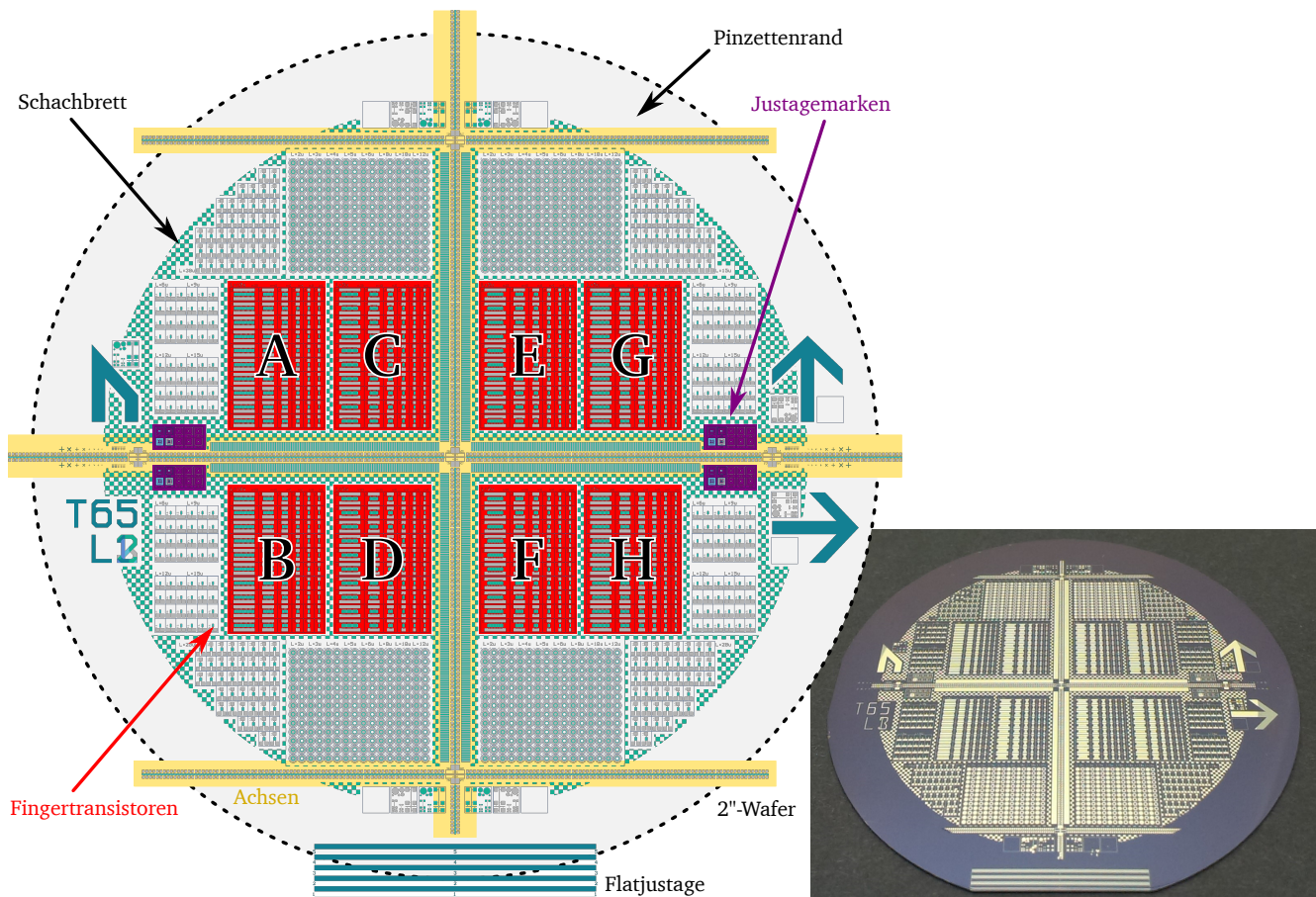


Abb. 35: Übersicht des Testchip-Layouts für 2"-Wafer mit acht Feldern mit Fingertransistoren (rot, A–H), Justagemarken (violett) und Pfeile, Flatjustage und Achsen (gelb) als makroskopische Orientierungshilfen.

Ebenen des Layouts

Entgegen dem in [6] beschriebenen Herstellungsprozess ohne strukturierten Katalysator umfasst das hier entworfene Layout drei Ebenen. Die Ebenen dienen im Einzelnen der Erzeugung der folgenden Strukturen:

1. Justagemarken und Gräben zur Abgrenzung der Transistoren gegeneinander
2. strukturierter Katalysator zum Wachsen der CNTs
3. Metallisierung der CNTs mit *Source/Drain*-Kontakten

Mit Ebene 1 werden für die folgenden Ebenen Justagemarken vorbereitet, die mittels Ätzung in das Siliziumdioxid strukturiert werden. Für die Justagemarken ist deswegen eine eigene Ebene im Layout vorgesehen, da die Schichtdicke des Katalysators sehr gering ist. Durch die geringe Schichtdicke ist das Justieren unter Gelblicht sehr schwierig, was eine direkte Justage der Ebene 3 auf Ebene 2 verhindert. Ein zusätzlicher Vorteil der eigenen Ebene ist zudem, dass auf diese Weise die Ätztiefe der Marken frei gewählt werden kann, sodass unter Gelblicht und durch den Lack hindurch ein optimaler Kontrast eingestellt werden kann.

Mit Ebene 2 wird der strukturierte Katalysator hergestellt. Die Positionierung der einzelnen „Katalysator-Inseln“ ist dabei so gewählt, dass die Inseln mittig in den durch die Gräben aus Ebene 1 definierten Gebieten zu liegen kommen. Mit Ebene 3 werden schließlich die *Source/Drain*-Kontakte hergestellt. Die Justage auf Ebene 1 ist dabei so ausgelegt, dass die Kontakte genau über die Katalysator-Inseln aus Ebene 2 positioniert werden.

Justagemarken

Wie die violetten Markierungen in Abb. 35 bereits zeigen, sind die Justagemarken in vier Gruppen zu zwei Reihen oberhalb und unterhalb der Mittelachse und zu zwei Spalten auf der rechten und linken Seite des Layouts angeordnet. Jedes Paar von Marken ist dabei nach dem gleichen Schema ausgeführt, das am Beispiel der Justage von Ebene 2 auf Ebene 1 in Abb. 36 erläutert wird: Die „untere“ Marke — also die Struktur, die sich bereits auf dem Wafer befindet — ist in Abb. 36a gezeigt; sie enthält in der Mitte ein Kreuz und am Rand nach innen gerichtete Kämme. Abb. 36b zeigt die „obere“ Marke — also die Struktur, die sich zunächst noch nur auf der Maske befindet —, die in der Mitte vier kleine Quadrate enthält, die das Kreuz aus der unteren Marke zu einem großen Quadrat ergänzen. Die in der oberen Marke enthaltenen, nach außen gerichteten Kämme greifen dabei exakt in die nach innen gerichteten Kämme der unteren Marke. Die Überlagerung der beiden Marken und damit ein optimales Ergebnis der Justage zeigt Abb. 36c.

Durch das Vorhandensein von vier Gruppen zusammen mit der Kamm-Struktur der Marken ist die Justage doppelt-redundant gegen Fehler in den lithographischen Prozessen abgesichert: Bei Ausfall einer Gruppe kann mit der jeweils anderen Reihe dennoch bequem eine Zwei-Punkt-Justage unter Verwendung der Kreuze und vier Quadrate durchgeführt werden, während im

Extremfall beim Ausfall von bis zu drei Gruppen auch mit einer einzigen Gruppe unter Verwendung der Kämme justiert werden kann. Zudem ist für jede einzelne Marke zusätzlich eine invertierte Ausführung vorgesehen worden, um eventuelle Änderungen in den lithographischen Prozessen kompensieren zu können.

Fingertransistoren

Jeder Transistor, der in einem der acht in Abb. 35 rot markierten Felder liegt, ist ein „Fingertransistor“ und grundlegend gleich aufgebaut. Ein exemplarischer Transistor ist in Abb. 37 als Ausschnitt aus dem Layout und als Schema gezeigt: Mit Ebene 1 des Layouts werden durch direkte Strukturübertragung mittels Ätzen Gräben (weiß) erzeugt, die die einzelnen Transistoren optisch gegeneinander abgrenzen und damit eine bessere Orientierung auf den Wafern ermöglichen. Durch die Gräben werden rechteckige Bereiche definiert, in die mittig mittels *Lift off*-Technik und Ebene 2 des Layouts Katalysator-Inseln (türkis) positioniert werden. Auf den Katalysator-Inseln wachsen während des CCVD-Schritts die CNTs. *Source/Drain*-Kontakte (grau) und damit eine Möglichkeit zur Kontaktierung der CNTs nach außen hin werden schließlich mit Ebene 3 realisiert. Jeder Kontakt besteht aus einem breiten Anschluss-Pad und schmalen Fingern, die über der Katalysator-Insel liegen und den eigentlichen Kontakt zu den CNTs herstellen, während die Pads selbst die Inseln nicht direkt berührt.

Wie auch in Abb. 37a gezeigt, sind bei allen Transistoren die beiden Anschluss-Pads $200\text{ }\mu\text{m}$ und die dünnen Finger $25\text{ }\mu\text{m}$ breit. Die Finger haben jeweils eine Gesamtlänge von $250\text{ }\mu\text{m}$, wobei sie sich in der Weite auf $200\text{ }\mu\text{m}$ überlappen. Die Katalysator-Inseln darunter haben eine Weite von $180\text{ }\mu\text{m}$, was hinsichtlich der Justage eine Toleranz von jeweils $10\text{ }\mu\text{m}$ nach oben und unten gestattet. Es wird also eine nominale Kanalweite pro Fingerpaar von $180\text{ }\mu\text{m}$ erreicht.

Das Layout enthält acht Felder mit Fingertransistoren, wobei über- und unterhalb der Mittelachse liegende Felder jeweils paarweise identisch sind. Die vier Gruppen unterscheiden sich nach

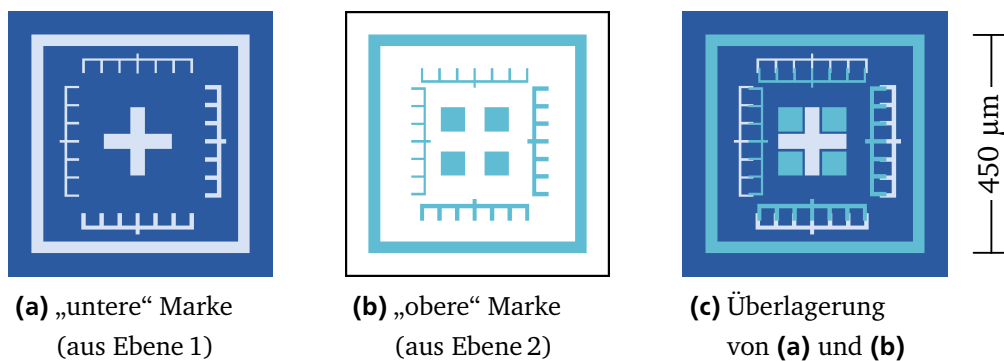


Abb. 36: Justagemarken des Testchip-Layouts: Die Justage erfolgt unter Verwendung zweier korrespondierenden Sätze bequem als Zwei-Punkt-Justage über die vier Quadrate und das Kreuz in der Mitte der Marken oder aufwendiger als Ein-Punkt-Justage über die Kamm-Strukturen am Rand der Marken.

der nominalen Länge der Kanäle $L_{\text{Kanal|nom}}$, wobei Längen von 2, 3, 5 und 10 μm vorgesehen sind. Jedes Feld hat sechs Spalten, die in zehn Zeilen jeweils identische Transistoren enthalten. Die Spalten unterschieden sich durch die Anzahl der Fingerpaare n , wobei Werte von 1, 2, 4, 8, 16 und 32 vorhanden sind. Eine Übersicht über die in den acht Feldern und sechs Spalten verfügbare Kanallängen und -weiten gibt Tab. 2, wobei jede Kombination aus Weite $W_{\text{Kanal|nom}}$ und Länge $L_{\text{Kanal|nom}}$ jeweils 20mal im Layout vorgesehen ist.

Die Verwendung des strukturierten Katalysators hat gegenüber dem Prozess mit unstrukturiertem Katalysator Vorteile. Zum einen bleibt das CNT-Wachstum aufgrund des katalytischen Wachstumsmechanismus auf die Katalysator-Inseln und die sie umgebenden in Abb. 37b skizzierten „Randbereiche“ beschränkt, womit ohne weitere Maßnahmen eine laterale Isolation der verschiedenen Bauelemente gegeneinander gegeben ist. Zum anderen können in den Randbereichen, wo der Untergrund das glatte thermische Oxid ist, präzisere AFM-Messungen durchgeführt werden als es auf dem vergleichsweise rauen Katalysator möglich ist.

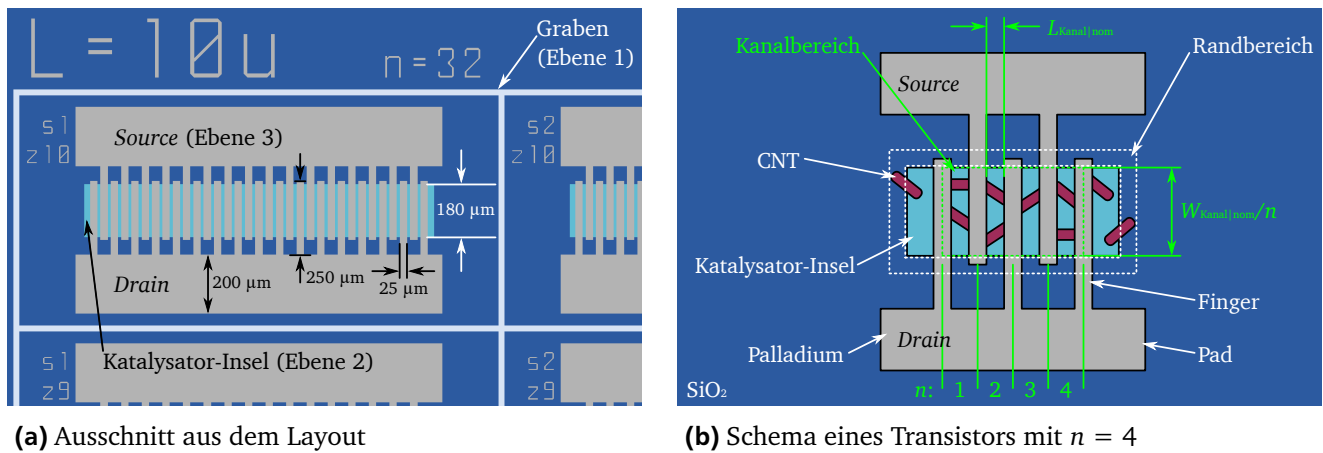


Abb. 37: Testchip-Layout für Fingertransistoren mit drei Ebenen für: in das Siliziumdioxid (blau) geätzte Gräben (weiß), strukturierten Katalysator (türkis) und Source/Drain-Kontakte (grau). Das CNT-Wachstum bleibt auf die Katalysator-Inseln beschränkt wodurch eine laterale Isolation zu den benachbarten Bauelementen erreicht wird. Die nominale Kanalweite ist über die Anzahl der Fingerpaare n geben. Die in (a) angegebenen Maße sind für alle Fingertransistoren unabhängig von Kanallänge und -weite gleich.

Tab. 2: In den acht Feldern und sechs Spalten verfügbare nominale Kanallängen und -weiten im Testchip-Layout (vgl. Abb. 35), wobei jede Kombination aus Weite und Länge jeweils 20mal im Layout vorgesehen ist.

Feld	A & B		C & D	E & F		G & H	
$L_{\text{Kanal nom}}$	10		3		2	5	μm
Spalte	32	16	8	4	2	1	Fingerpaare (n)
$W_{\text{Kanal nom}}$	5760	2880	1440	720	360	180	μm

6.2 Prozessablauf

In Abb. 38 ist der Prozessablauf zur Herstellung von CNTFETs als Blockdiagramm graphisch dargestellt wie er ausgehend von der Vorgängerarbeit [6] entwickelt wurde. Sämtliche Arbeitsschritte werden dabei im Reinraum und den Laboren des IHTN durchgeführt. Der Prozessablauf gliedert sich in sechs Blöcke, die auf den folgenden Seiten näher erläutert werden:

1. Vorbereitung der Substrate
2. Herstellung von Justage-Marken
3. Abscheidung und Strukturierung des Katalysators
4. Cluster-Bildung und CNT-Wachstum
5. Herstellung von *Source/Drain*-Kontakten und *Backgate*
6. Charakterisierung der gefertigten Bauelemente

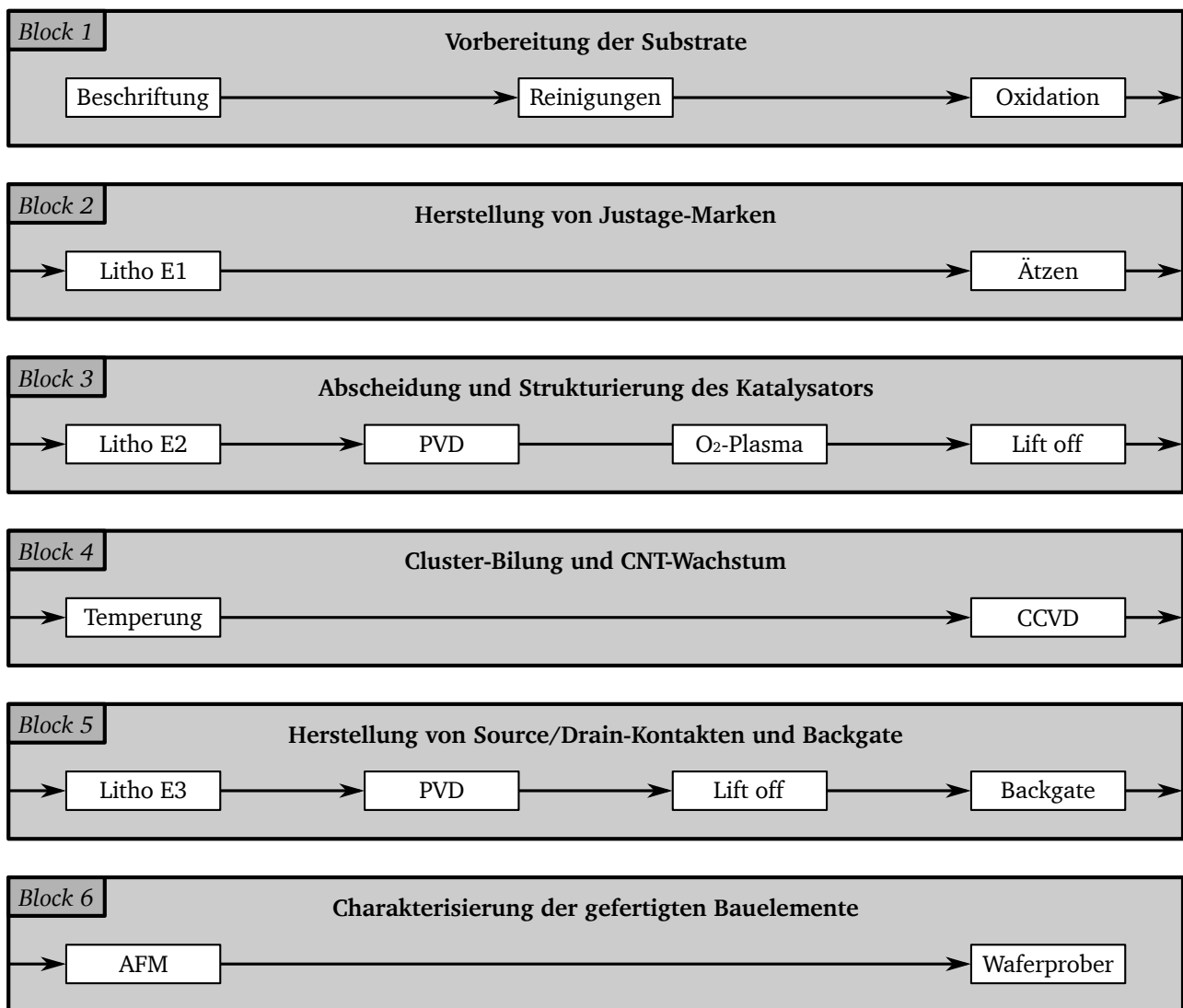


Abb. 38: Prozessablauf zur Herstellung von CNTFETs wie er aufbauend auf [6] entwickelt wurde.

Für den Block „Cluster-Bildung und CNT-Wachstum“, der die beiden zentralen Prozessschritte der Herstellung umfasst, werden in Kapitel 7 (Seite 67) die Details der Prozessführung ausführlicher dargestellt. Hinsichtlich der Anlagentechnik und der verschiedenen Prozessschritte im Allgemeinen und der *Lift off*-Technik im Besonderen sei auf Kapitel 5 (Seite 33) bzw. Kapitel 5.5 (Seite 40) verwiesen. Ein exemplarischer Chargenbegleitschein mit den Basisparametern der Herstellung ist ab Seite 133 in den Anhang gestellt.

Block 1: Vorbereitung der Substrate

Als Substrate werden Bor-dotierte 2"-Wafer mit einem spezifischen Widerstand von 0,01 bis 0,02 $\Omega\cdot\text{cm}$ verwendet, die mittels des Czochralski-Verfahrens in (100)-Richtung gezogen wurden. Nach dem Beschriften erfolgt die Reinigung der Wafer: Grobe Partikel werden in einer Kaskade mit deionisiertem Wasser (DI-Wasser) abgespült, um keine makroskopischen Verunreinigungen in die folgenden Bäder zu verschleppen. Die Entfernung organischer Verunreinigungen geschieht mit einer Mischung aus Salpeter- und Schwefelsäure bei zunächst 110 °C und schließlich bei Raumtemperatur für 10 min bzw. 5 min und anschließend mit einer Mischung aus Wasserstoffperoxid und Schwefelsäure für 15 min („*piranha clean*“). Das natürliche Oxid der Silizium-Wafer wird danach in einer wässrigen 0,5 %igen Lösung der Flusssäure entfernt („*HF-Dip*“). Zur Reinigung von feinen Partikeln und zur Erzeugung einer passivierenden Schicht wird für die Dauer von 10 min eine auf 70 °C temperierte Mischung von DI-Wasser, Wasserstoffperoxid und Ammoniumhydroxid genutzt („*standard clean 1*“ oder auch „*RCA clean 1*“). Abschließend werden metallische Kontaminationen in 2 min mit einer stark verdünnten wässrigen Lösung der Salzsäure entfernt, deren pH-Wert auf kleiner 2 eingestellt wurde.

Nach der Reinigung wird ein weiterer *HF-Dip* und eine Oxidation durchgeführt, um ein *Gate-Oxid* mit einer Dicke von 100 nm zu wachsen, wobei Abb. 39 das Ergebnis skizziert. Die Oxidation erfolgt in trockener Sauerstoff-Atmosphäre bei 1000 °C für 120 min in einem Horizontal-Rohr-Ofen des Herstellers *Thermoco*. Die Prozesskontrolle wird wie in Kapitel 4.1 (Seite 25) beschrieben mittels Ellipsometrie durchgeführt.

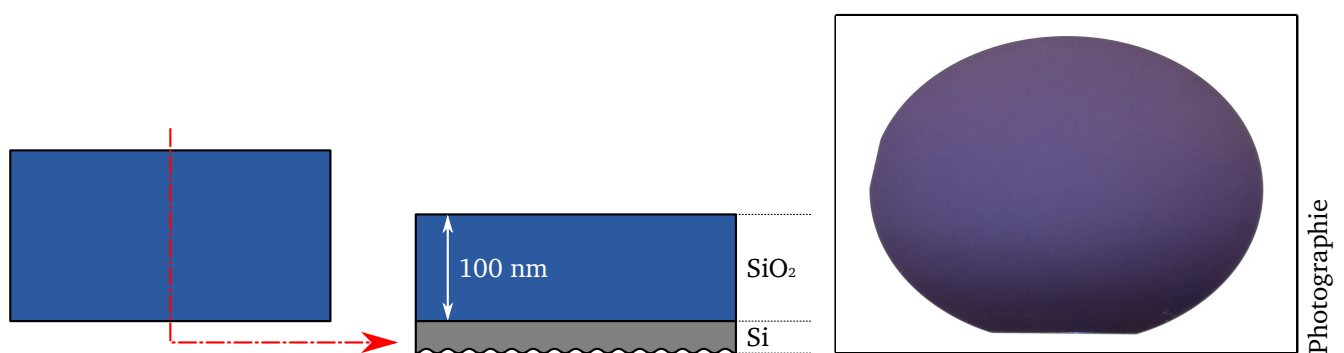


Abb. 39: Gereinigtes und oxidiertes 2"-Substrat.

Block 2: Herstellung von Justage-Marken

Die Justagemarken werden unter Verwendung von Ebene 1 des Testchip-Layouts mittels Strukturübertragung durch nasschemisches Ätzen im Oxid erzeugt. Diese Marken dienen der Justage der beiden folgenden Ebenen des Layouts, die den strukturierten Katalysator und die *Source/Drain*-Kontakte enthalten. Gleichzeitig werden mit dieser Ebene, wie in Abb. 40 gezeigt, Gräben erzeugt, die die Transistoren gegeneinander abgrenzen und einer besseren Orientierung während der verschiedenen Messungen auf den Wafern dienen.

Für die Strukturübertragung werden die oxidierten Substrate zunächst mit Hilfe einer Lack-schleuder des Herstellers *Convac* mit einer $1,4\text{ }\mu\text{m}$ dicken Lackschicht des Lacks *AZ 5214E* von *MicroChemicals* versehen. Nach der Trocknung des Lacks auf einer Heizplatte erfolgt die Belichtung für 25 s im Kontakt-Modus mittels des Belichters *MA 56M* des Herstellers *Karl Süss*, der eine Quecksilberdampfampe nutzt. Die Entwicklung wird anschließend in *AZ 726 MIF* von *MicroChemicals* für 1 min durchgeführt.

Zum Ätzen bis in eine Tiefe von 80 nm wird gepufferte Oxid-Ätze (BOE) verwendet, um eine Delamination des Lacks zu verhindern und möglichst scharfe Kanten zu erhalten. Die Tiefe von 80 nm wurde gewählt, da so noch eine die Silizium-Oberfläche schützende Oxidschicht in den Gräben vorhanden ist, gleichzeitig aber auch unter Gelblicht und durch den Lack hindurch ein ausreichender Kontrast zum Justieren besteht. Nach Ablacken in *n*-Ethyl-2-pyrrolidon (NEP) bei 70°C wird eine Kontrolle der Grabentiefe mittels AFM durchgeführt. Abschließend werden die Wafer für 15 min in einer *piranha clean*-Lösung gereinigt, um eventuelle Reste des Lacks zu entfernen.

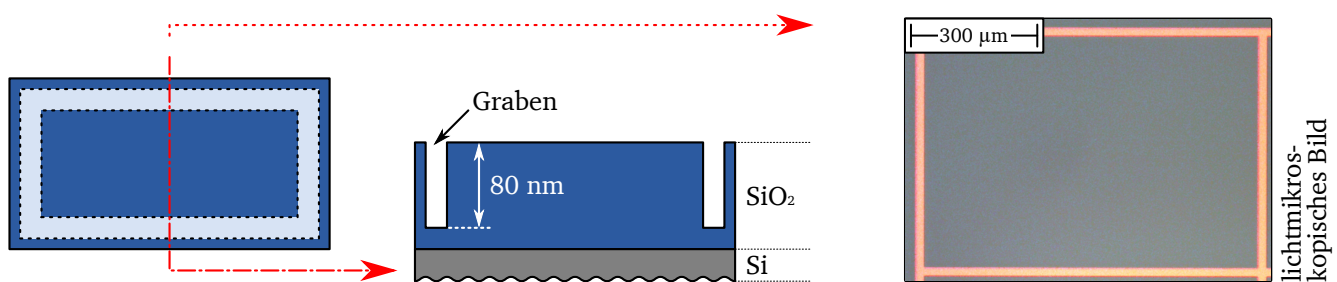


Abb. 40: Geätzte Gräben zur Abgrenzung der Transistoren und Justagemarken (nicht gezeigt).

Block 3: Abscheidung und Strukturierung des Katalysators

Zur Herstellung des strukturierten Katalysators wird das PVD-Verfahren der Elektronenstrahlverdampfung in Verbindung mit der *Lift off*-Technik genutzt. Zur Erzeugung der nötigen Lackstruktur mit Unterschnitten wird wie in Block 2 eine $1,4\text{ }\mu\text{m}$ dicke Lackschicht aufgebracht und im Kontakt-Modus belichtet. Als Dauer der Belichtung werden 5 s gewählt, sodass eine Unterbelichtung erreicht wird. Die thermische Umkehrung des Lacks erfolgt bei 125°C für 90 s auf einer Heizplatte. Nach der Flutbelichtung mit einer Dauer von 2 min wird die Entwicklung in

AZ 726 MIF in 3 min durchgeführt. Nach der Entwicklung wird ein *HF-Dip* durchgeführt, um eine saubere und trockene Oberfläche sicherzustellen. Danach werden in einem gemeinsamen Prozessschritt in der Bedampfungsanlage *BPU 100* des Herstellers *Balzers* zuerst 5 nm Aluminium und schließlich 1 nm Nickel aufgedampft. Im Anschluss an die Bedampfung und die Sauerstoff-Plasmabehandlung wird der *Lift off* in NEP bei 70 °C mit Ultraschall-Unterstützung mit dem in Abb. 41 skizzierten Ergebnis durchgeführt.

Bei der Sauerstoff-Plasmabehandlung, die zwischen der Bedampfung und dem *Lift off* durchgeführt wird, handelt sich um einen trockenchemischen Plasma-Ätzprozess (CDE). Dieser Prozess hat bei einem Druck von 200 mTorr eine Dauer von 5 min und wurde am IHTN ursprünglich in ähnlicher Form zur Entfernung von Lackresten nach der Strukturübertragung durch Ätzen oder nach Ionenimplantationen verwendet („Lackveraschen“, auch „*O₂-Dip*“). Die Wafer befinden sich während des Prozesses auf der Erdpotential-Elektrode des Trockenätzers (Eigenbau des IHTN mit einer Plasmaquelle des Herstellers *Hüttinger*), während über die andere Elektrode eine hochfrequente Wechselspannung mit einer Leistung von 100 W eingekoppelt wird.

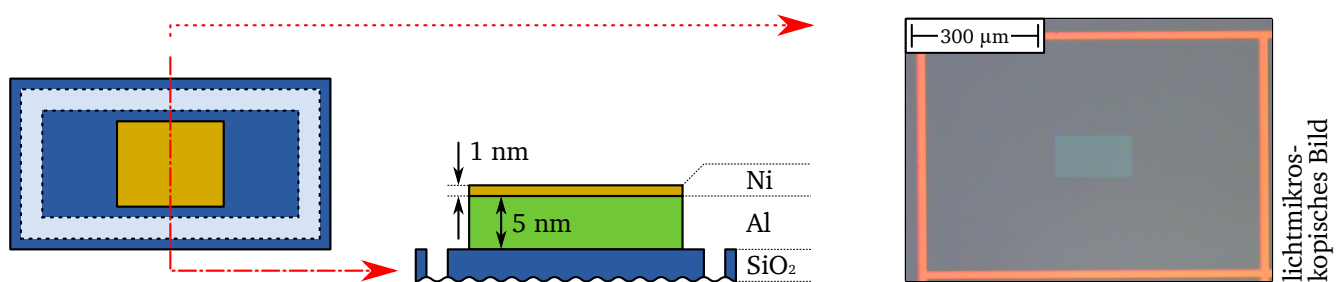


Abb. 41: Mittels *Lift off*-Technik und Elektronenstrahlverdampfung strukturierter Katalysator.

Block 4: Cluster-Bildung und CNT-Wachstum

Die zwei zentralen Prozessschritte sind die Cluster-Bildung und das CNT-Wachstum, die beide in der CVD-Anlage *AVM 1200* des Herstellers *Applied Materials* durchgeführt werden. Die Cluster-Bildung erfolgt durch eine Temperung in Inertgas-Atmosphäre, wobei der metallische Doppelstapel — wie auch in Abb. 42 schematisch gezeigt — zu einem dielektrischen Aluminiumoxid und diskreten Nickel-Clustern transformiert wird.

Anschließend werden, wie in Abb. 43 skizziert, in zunächst 10 min mit Hilfe von Methan und der zuvor gebildeten Nickel-Cluster entsprechend Gl. 43 CNTs gewachsen. Wasserstoff wird neben Methan in die Anlage eingeleitet, da er selbst ein Reaktionsprodukt von Gl. 43 ist und damit die Reaktionsgeschwindigkeit reduziert. Auf diese Weise wird sichergestellt, dass die Pyrolyse des Methan nicht unkontrolliert in der Gasphase stattfinden kann.



Eine Prozesskontrolle der Cluster-Bildung und des CNT-Wachstums kann mittels AFM erfolgen.

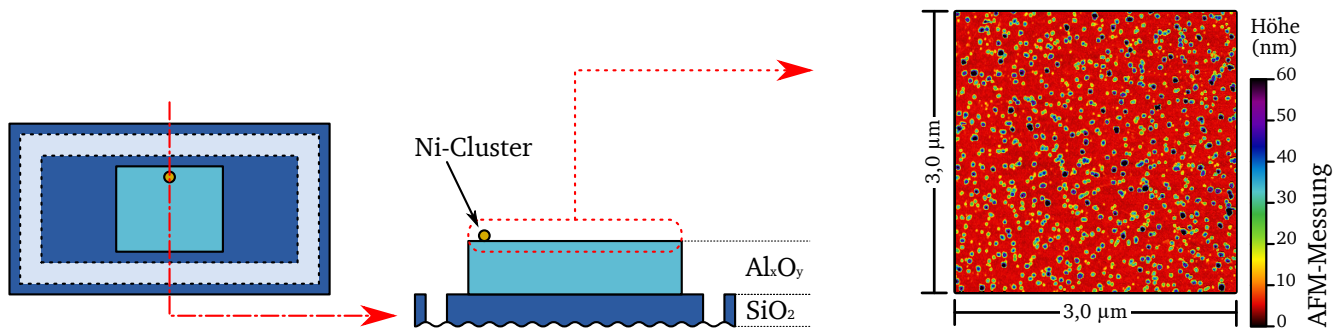


Abb. 42: Getemperter Katalysator mit Nickel-Clustern.

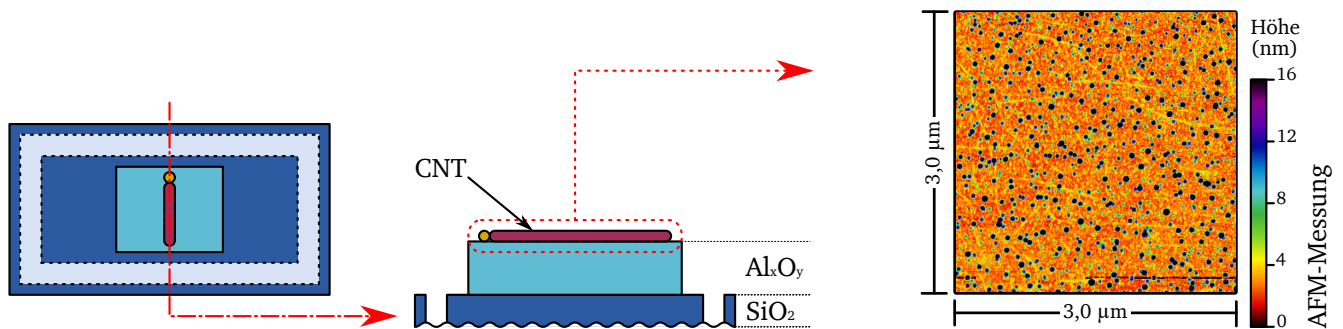


Abb. 43: Gemäß Gl. 43 gewachsene CNT.

Block 5: Herstellung von *Source/Drain*-Kontakten und *Backgate*

Als letzter Block vor der abschließenden Auswertung folgt das Aufbringen von *Source/Drain*-Kontakten aus Palladium und die Herstellung des *Backgates*. Die Strukturierung der *Source/Drain*-Kontakte erfolgt mittels *Lift off*-Technik und Ebene 3 des Testchip-Layouts und durch Elektronenstrahlverdampfung von Palladium. Die lithographischen Arbeitsschritte im Rahmen der *Lift off*-Technik erfolgen dabei mit den gleichen Materialien und Parametern wie bei der Erzeugung des strukturierten Katalysators in Block 3. Für die Herstellung des *Backgates* wird das thermisch gewachsene Oxid auf der Rückseite der Wafer selektiv mit BOE entfernt. Einen auf diese Weise prozessierten Transistor zeigt Abb. 44.

Palladium wird wegen seiner günstigen Austrittsarbeit als Kontaktmaterial verwendet [74]. Da Palladium aber sehr weich ist und zudem nur schlecht auf dem Siliziumdioxid-Untergrund haftet, werden die Kontakte als Doppellagen-System ausgeführt. (Die problematischen Eigenschaften von reinen Palladium-Kontakten werden im Anhang ab Seite 135 eingehender beschrieben.) Wie ebenfalls in Abb. 44 skizziert, werden zunächst 10 nm Palladium abgeschieden, um den gewünschten Kontakt zur CNT herzustellen; dann werden im gleichen Schritt der Bedampfung 100 nm Nickel aufgedampft („*Capping*“), um robuste Kontakte für die elektrische Charakterisierung zu erhalten. In direktem Kontakt mit der CNT ist dabei nur die Palladium-Schicht, d. h. das Palladium wird an der Kante nicht von Nickel überlappt, da bei der Strukturierung mittels *Lift off* durch den Unterschnitt während der Bedampfung eine Abschattung vorliegt und somit die Schichten nach oben hin schmaler werden [75].

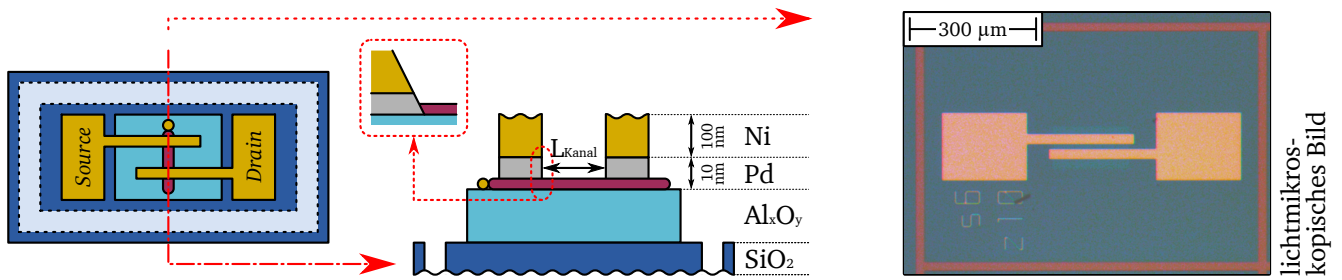


Abb. 44: Vollständig prozessierter Transistor mit Palladium-Kontakten und Nickel-Capping.

Block 6: Charakterisierung der gefertigten Bauelemente

Die abschließende Charakterisierung der Bauelemente erfolgt mittels AFM im *tapping mode* und elektrischer Messungen mit einem Waferprober. Die verwendeten Messgeräte sind in Kapitel 4.2 (Seite 28) bzw. in Kapitel 4.3 (Seite 32) beschrieben.

Bei den AFM-Messungen wird üblicherweise ein quadratischer Messbereich mit einer Kantenlänge von nicht mehr als $5\text{ }\mu\text{m}$ gewählt. Bei 1024×1024 im gesamten Messbereich aufgenommenen Datenpunkten ergibt sich damit eine digitale laterale Auflösung von knapp 5 nm pro Datenpunkt. Die verwendeten Messspitzen haben einen Radius von 15 nm , was dann in Abhängigkeit von der auf der Probe vorliegenden Topologie zu einer tatsächlichen lateralen Auflösung von mehr als 5 nm führt.

Für die elektrische Charakterisierung wird eine Kontaktierung und Benennung der Größen gemäß Abb. 45 gewählt. Da es sich hier zunächst um Bauelemente mit p-MOS-Verhalten handelt, wird die Spannung V_{DS} als negatives Potential an den Drain-Kontakt angelegt, während der Source-Kontakt auf 0 V gegen Erdpotential gehalten wird. Sofern nicht explizit anders angegeben wird in dieser Arbeit ein V_{DS} von -500 mV gewählt.

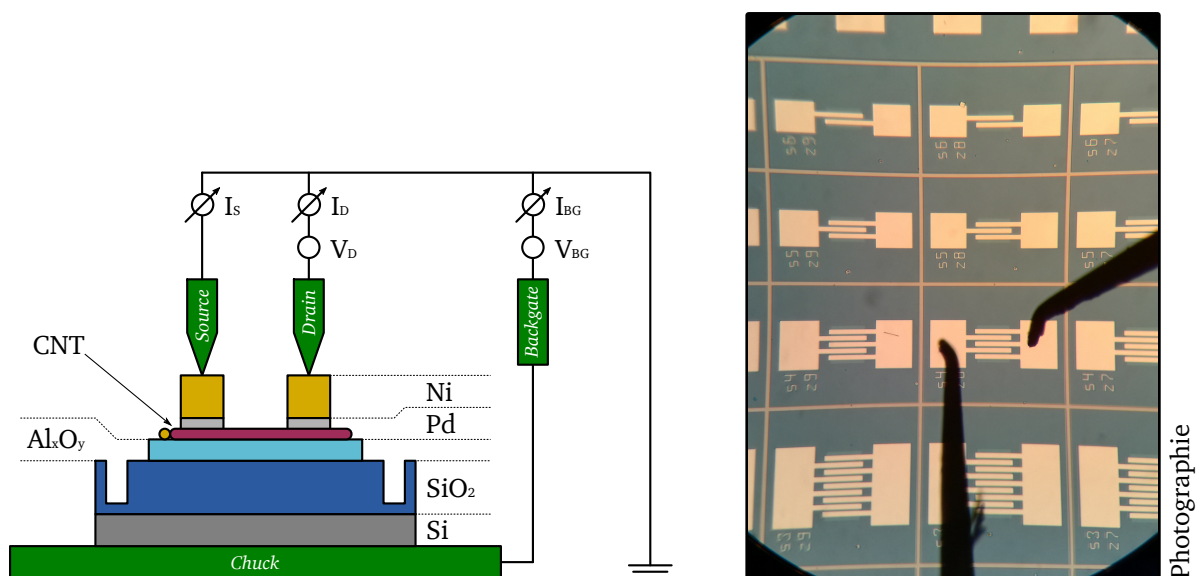


Abb. 45: Aufbau zur elektrischen Charakterisierung der prozessierten Transistoren.

7 Topologische Untersuchung von Cluster-Bildung und CNT-Wachstum

In diesem Kapitel soll ermittelt werden, inwiefern durch die verschiedenen Prozessparameter die Cluster-Bildung und das CNT-Wachstum gezielt beeinflusst werden kann. Der Fokus liegt dabei auf der Prozessführung während der Temperung und während des CCVD-Schritts, es wird aber auch die Sauerstoff-Plasmabehandlung des Katalysators betrachtet, die in dieser Arbeit bei der Herstellung von CNTFETs obligatorisch ist. Ausgehend vom Prozess der Vorgängerarbeit [6] werden verschiedene Variationen der Prozessführung während der Temperung und des CCVD-Schritts vorgenommen und im Anschluss die gebildeten Cluster und gewachsenen CNTs mittels AFM untersucht. Die Bewertung der Messergebnisse erfolgt dabei unter der Zielsetzung einen Prozess zu finden, mit dessen Hilfe eine hohe Anzahl möglichst langer und möglichst dünner CNTs *in situ* gewachsen werden kann.

Sämtliche hier gezeigten Proben werden wie in Kapitel 6.2 (Seite 60) beschrieben hergestellt, wobei die Prozessierung nach dem CCVD-Schritt beendet wird. Es wird also zunächst auf die Herstellung von *Source/Drain*-Kontakten verzichtet, weswegen auch die Erzeugung der Justage-Marken mit Ebene 1 des Testchip-Layouts entfallen kann. Die Proben haben also folgende Eckdaten:

- Bor-dotierte Silizium-Substrate
- 100 nm thermisch gewachsenes Trocken-Oxid
- strukturierter Katalysator bestehend aus 5 nm Aluminium und 1 nm Nickel
- Sauerstoff-Plasmabehandlung des Katalysators

7.1 Beschreibung des CNT-Wachstumsmechanismus

Das katalytische Wachstum von CNTs auf Basis von nur wenige Nanometer großen metallischen Partikeln („Cluster“), unterscheidet zwischen dem *base growth*- und dem *tip growth*-Modus. Der *base growth*-Modus beschreibt ein Wachstum, bei dem die CNT aus dem Cluster heraus wächst, wobei der Cluster seine Position beibehält, wogegen beim *tip growth*-Modus die wachsende CNT den Cluster vor sich herschiebt. Welcher dieser beiden Modi vorliegt, wird dabei nach GOHIER ET AL. von der Größe der Cluster bestimmt: für kleine Cluster liegt *base growth* vor, für große

Cluster *tip growth*. Als Grenzwert für die Unterscheidung der Cluster-Größe wurde für einen Plasma-unterstützten CCVD-Prozess und für verschiedene Katalysatoren 5 nm ermittelt [30].

Parallel zur Aufteilung in *base growth* und *tip growth* wird noch der *vapor-liquid-solid*- und der *vapor-solid-solid*-Transport unterschieden. Der *vapor-liquid-solid*-Transport geht von einer Absorption des Kohlenstoffs im Cluster und einer Volumendiffusion aus, wobei als Zwischenstufe ein flüssiges, metastabiles Karbid mit dem Material des Clusters gebildet wird. Der *vapor-solid-solid*-Transport basiert auf einer Adsorption des Kohlenstoffs am Cluster und einer Oberflächendiffusion, wobei der Kohlenstoff in seiner atomaren Form verbleibt und keine Verbindung mit dem Material des Clusters eingeht. Die Frage, welcher dieser beiden Transportarten zutrifft, ist noch Gegenstand aktueller Forschung. TESSONNIER UND SU kommen in ihrem Review zu dem Schluss, dass der in der Forschung historisch ältere *vapor-liquid-solid*-Transport überholt sei; sie führen die Berechnung von HOFMANN ET AL. an, die für die Oberflächendiffusion von Kohlenstoff auf einem Cluster eine niedrigere Aktivierungsenergie ergeben als für die Volumendiffusion [76, 77]. Die bereits zitierte Gruppe um GOHIER geht hingegen davon aus, dass beide Transportarten möglich sind und macht diesen Umstand zudem für den von ihnen festgestellten Wechsel von *base growth* zu *tip growth* für größere Cluster verantwortlich.

SEIDEL ET AL. haben unter anderem einen Prozess untersucht, der Parallelen zu dem in dieser Arbeit genutzten Prozess aufweist (ebenfalls Aluminium und Nickel als Katalysator und ähnliche Parameter für die Temperatur und den CCVD-Schritt). Das Wachstum läuft, ihrer Argumentation folgend, im *base growth*-Modus und mit *vapor-solid-solid*-Transport über einen vierstufigen Prozess ab, der in Abb. 46 skizziert ist [31]:

a) Beginn der Belegung:

Der Prozess umfasst zunächst die Dissoziation bzw. Pyrolyse eines Kohlenstoff-haltigen Gases, zumeist einer Kohlenwasserstoff-Verbindung, an der Oberfläche des Clusters. Der Kohlenstoff wird an der Oberfläche des Clusters adsorbiert, während der Wasserstoff und eventuelle weitere Bestandteile des Gases über die Gasphase abtransportiert werden.

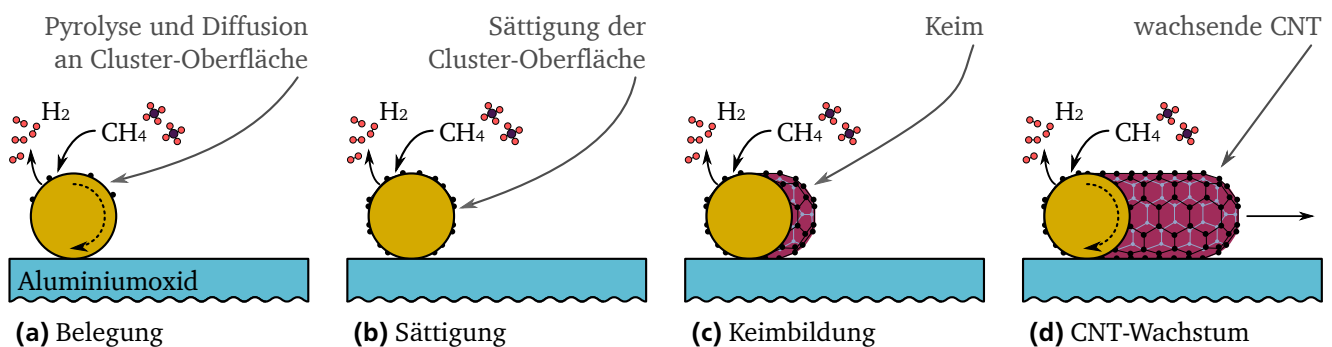


Abb. 46: Vierteiliger Wachstumsmechanismus von CNTs auf Basis des *base growth*-Modus und des *vapor-solid-solid*-Transports nach [31] unter Verwendung von Methan als Kohlenstoff-Quelle.

b) Sättigung der Oberfläche:

Weiterer Kohlenstoff adsorbiert am Cluster und verteilt sich über Oberflächendiffusion bis die Oberfläche gesättigt ist.

c) Bildung eines Keims:

Der Kohlenstoff bildet dabei eine Hülle um den Cluster herum in Form einer Fulleren-artigen, pentagonalen und hexagonalen Struktur aus. Dabei ist es unwahrscheinlich, dass diese Hülle die Oberflächenkrümmung des Clusters vollständig nachbildet, sodass zwischen Hülle und Cluster eine mechanische Spannung entsteht. Diese Spannung relaxiert, indem sich die Hülle leicht von der Oberfläche des Clusters abhebt — und somit den Keim für eine CNT bildet. Die Form des Keims bestimmt dabei den Durchmesser und die Chiralität der CNT, weswegen diese Stufe für den Wachstumsmechanismus und für die späteren Eigenschaften der CNT essentiell ist [78].

d) Wachstum einer CNT:

Im weiteren Verlauf wird weiterer Kohlenstoff über Pyrolyse aus der Gasphase am Cluster adsorbiert, über Oberflächendiffusion zur Grenzfläche mit der sich ausbildenden CNT transportiert und dort angelagert, während der Wasserstoff oder weitere Bestandteile des Gases weiterhin zurück in die Gasphase desorbieren und über diese abtransportiert werden.

Unabhängig von den oben kurz angerissenen theoretischen Details lautet die grundlegende Aussage aller Theorien zum Wachstumsmechanismus von CNTs, dass der Durchmesser einer CNT zur Größe des zugehörigen Clusters proportional ist. Im Hinblick auf die Herstellung von CNTFETs auf Basis von *in situ* gewachsenen CNTs besteht damit ein Zusammenhang zwischen der Größe der Cluster und dem elektrischen Verhalten der Transistoren. Da für CNTFETs halbleitende CNTs mit möglichst großer Bandlücke, d. h. CNTs mit möglichst geringem Durchmesser, benötigt werden, muss die Prozessführung während der Cluster-Bildung das Ziel haben, möglichst kleine Cluster zu erzeugen.

7.2 Beschreibung des Basisprozesses

Der Basisprozess entspricht der Prozessführung der Vorgängerarbeit [6] und stellt den Ausgangspunkt für die verschiedenen Variationen der Prozessführung dar, die in den folgenden Unterkapiteln diskutiert werden; er wird vollständig in der CVD-Anlage und mit Stickstoff als Inertgas durchgeführt. Dieser Prozess ist in Abb. 47 skizziert und umfasst die folgenden Schritte: Nach dem Beschicken der CVD-Anlage wird die Reaktionskammer mit zunächst 16 L/min und danach 8 L/min Stickstoff für 3 min bzw. 7 min gespült („Spülen“). Nach dem Spülen beginnt die

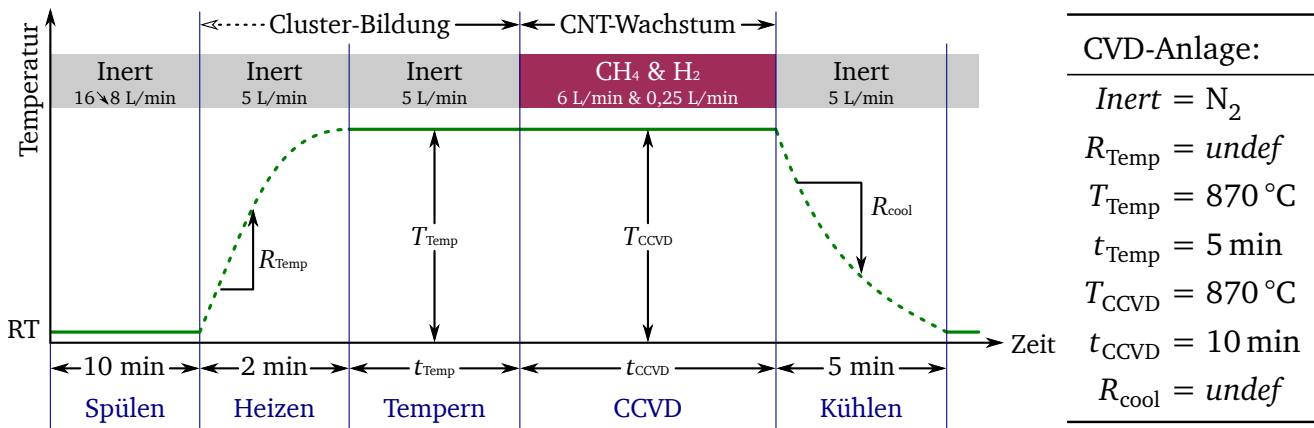


Abb. 47: Basisprozess nach [6] unter Verwendung der CVD-Anlage, wobei während der Temperung die Cluster gebildet und während des CCVD-Schritts die CNTs gewachsen werden. Hierbei sind: T_{Temp} und t_{Temp} die Temperatur bzw. die Dauer der Cluster-Bildung; T_{CCVD} und t_{CCVD} die Temperatur bzw. die Dauer des CNT-Wachstums; R_{Temp} die Heizrate der Cluster-Bildung; $Inert$ das verwendete Inertgas; R_{cool} die Kühlrate nach dem CNT-Wachstum und „RT“ die Raumtemperatur. Die Schritte „Heizen“ und „Kühlen“ erfolgen anlagentechnisch bedingt mit jeweils undefinierten Raten.

Cluster-Bildung mit einer Aufheizphase in Stickstoff-Atmosphäre, die rund 2 min dauert („Heizen“). Nach Erreichen der Cluster-Bildungs-Temperatur T_{Temp} von $870\text{ }^{\circ}\text{C}$, werden die Proben für eine Dauer t_{Temp} von 5 min getempert („Tempern“). Während dieser beiden Schritte läuft die Umwandlung des Katalysators ab, sobald die Erweichungstemperatur von Aluminium überschritten wurde, d. h. oberhalb von etwa $600\text{ }^{\circ}\text{C}$. (Durch den offenen, gestrichelten Pfeil soll gekennzeichnet werden, dass die Cluster-Bildung nicht bereits bei Raumtemperatur einsetzt.) Danach wird die Gasversorgung von Stickstoff auf die Methan/Wasserstoff-Mischung umgestellt, um bei der CNT-Wachstums-Temperatur T_{CCVD} von ebenfalls $870\text{ }^{\circ}\text{C}$ für die Dauer t_{CCVD} von 10 min CNTs zu wachsen („CCVD“). Nach dem CNT-Wachstum wird die Gasversorgung wieder auf Stickstoff umgeschaltet und der Heizer der Anlage abgeschaltet, worauf die Proben abkühlen und nach rund 5 min entnommen werden („Kühlen“). Der Stickstoff-Fluss beträgt während der Cluster-Bildung und während des Kühlens jeweils 5 L/min; die Methan/Wasserstoff-Mischung setzt sich aus 6 L/min Methan und 0,25 L/min Wasserstoff zusammen. Zu beachten ist hier, dass aufgrund der Konstruktion der CVD-Anlage die Aufheiz- und die Kühlphase mit jeweils undefinierter Rate ablaufen.

7.3 Einfluss der Sauerstoff-Plasmabehandlung des Katalysators

Die hier und bereits in der Vorgängerarbeit genutzte Sauerstoff-Plasmabehandlung ist, wie bereits in Kapitel 5.11 (Seite 52) beschrieben, ursprünglich ein Reinigungsschritt zur Entfernung von Lackresten. Dieser trockenchemische Plasma-Ätzprozess (CDE) wird bei einem Druck von 200 mTorr durchgeführt und hat eine Dauer von 5 min. Der Wafer befindet sich während des

Prozesses auf der Erdpotential-Elektrode des Trockenätzers, während über die andere Elektrode eine hochfrequente Wechselspannung mit einer Leistung von lediglich 100 W eingekoppelt wird. Nach der Sauerstoff-Plasmabehandlung wird die Cluster-Bildung und das CNT-Wachstum mit Hilfe des Basisprozesses gemäß Abb. 47 durchgeführt.

Obwohl der Katalysator nur wenige Nanometer dick ist, ist dennoch nicht von vornherein davon auszugehen, dass eine solche Sauerstoff-Plasmabehandlung, die derart wenig Leistung hat und so kurz ist, überhaupt einen nennenswerten Einfluss auf die Cluster-Bildung und das Wachstum von CNTs hat. Dass ein solcher Einfluss vorhanden und noch dazu durchaus ausgeprägt ist, zeigt Abb. 48 anhand zweier Proben, von denen nur eine die Sauerstoff-Plasmabehandlung erfahren hat, die aber ansonsten vollständig gleich prozessiert wurden. Werden die AFM-Aufnahmen der Plasma-behandelten und der unbehandelten Probe verglichen, ist deutlich zu erkennen, dass die Anzahl an CNTs auf der behandelten Probe höher ist. Außerdem zeigt sich, dass der CNT-Durchmesser mit 1 bis 2 nm gegenüber rund 3 nm auf der unbehandelten Probe kleiner ist. Die Cluster-Bildung und das CNT-Wachstum werden durch die Sauerstoff-Plasmabehandlung also positiv beeinflusst. Es muss allerdings auch festgehalten werden, dass die Sauerstoff-Plasmabehandlung die Ergebnisse verbessert, aber nicht grundlegend bedingt: Auch ohne Sauerstoff-Plasmabehandlung werden Cluster gebildet und CNTs gewachsen; nur eben in geringerer Anzahl und mit größeren Durchmessern.

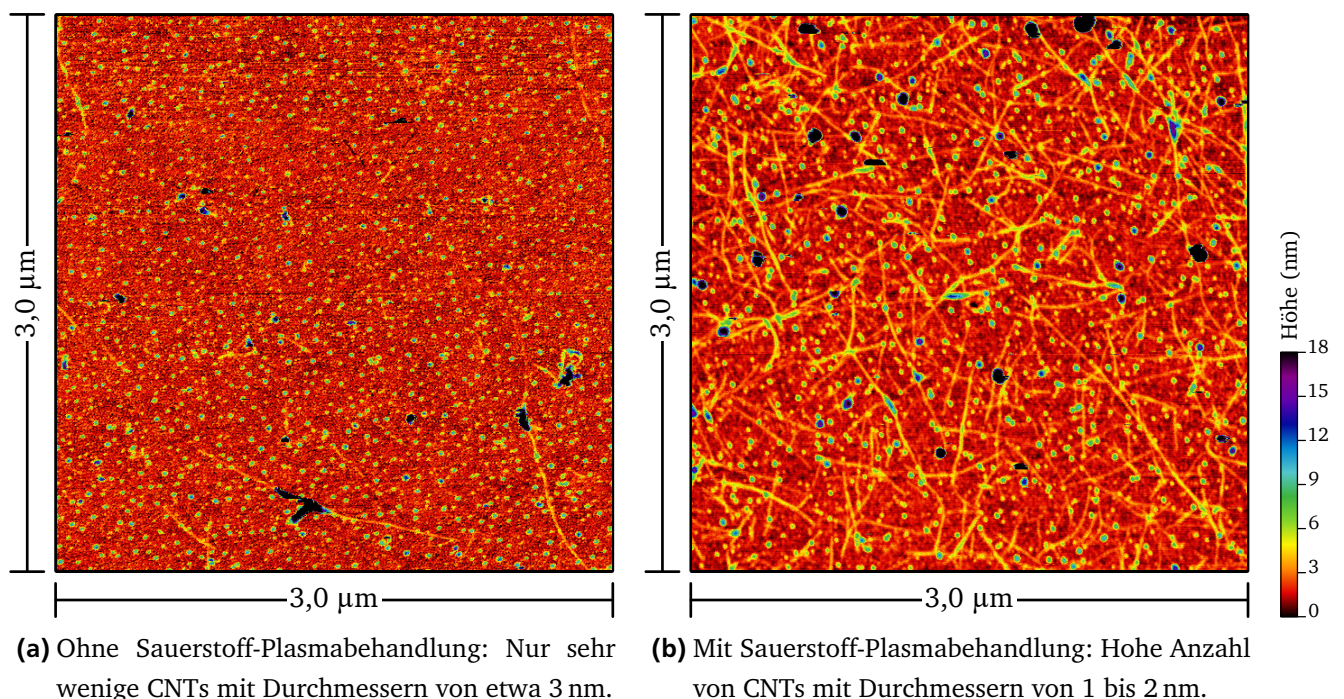


Abb. 48: AFM-Messungen auf Proben, die mit Ausnahme der Sauerstoff-Plasmabehandlung ansonsten identisch mit dem Basisprozess gemäß Abb. 47 hergestellt wurden. Der Einfluss der Sauerstoff-Plasmabehandlung ist klar zu erkennen: auf der behandelten Probe ist die Anzahl der CNTs deutlich höher; außerdem sind die gewachsenen CNTs dünner.

7.4 Einfluss der CNT-Wachstumsdauer

Als erster eigentlicher Parameter der Prozessführung soll die CNT-Wachstumsdauer t_{CCVD} betrachtet werden. Der dafür genutzte Prozess ist in Abb. 49 dargestellt, wobei Werte von 10 min und 20 min für t_{CCVD} gewählt wurden. Da die Wachstumsdauer t_{CCVD} von allen anderen Parametern des Basisprozesses unabhängig ist, kann sie in einfacher Weise direkt, d. h. ohne weitere Änderungen am Prozess, variiert werden, indem die Dauer des CCVD-Schritts verlängert wird.

Generell kann dabei davon ausgegangen werden, dass die Menge an Kohlenstoff, die für das Wachstum vom CNTs insgesamt zur Verfügung gestellt wird, direkt mit der Wachstumsdauer t_{CCVD} skaliert. Bei einer Verlängerung des CCVD-Schritts kann demnach für das CNT-Wachstum hinsichtlich der Anzahl und Länge der gewachsenen CNTs ein Ergebnis erwartet werden, dass innerhalb der drei folgenden Grenzfälle liegt:

1. Längere CNTs bei gleicher Menge; das hieße, dass der Einbau von Kohlenstoff in bereits wachsende CNTs gegenüber der Bildung eines neuen Keims bevorzugt wird.
2. Mehr CNTs bei gleichbleibender Länge; das hieße, dass das Wachstum ab einer gewissen Länge gehemmt wird.
3. Keine Beeinflussung der Menge oder der Länge; das hieße, dass der Wachstumsmechanismus an einem gewissen Punkt in eine Sättigung läuft, ab der weiterer Kohlenstoff durch die Cluster nicht weiter „verarbeitet“ werden kann.

Abb. 50 zeigt zwei AFM-Messungen, die auf Proben mit jeweils einer Wachstumsdauer von 10 min bzw. 20 min im Randbereich der Katalysator-Inseln durchgeführt wurden. Für eine Wachstumsdauer von 10 min können 14 CNTs im Messbereich beobachtet werden, deren durchschnittliche Länge $0,5 \mu\text{m}$ beträgt, wobei eine dieser CNTs eine überdurchschnittliche Länge von

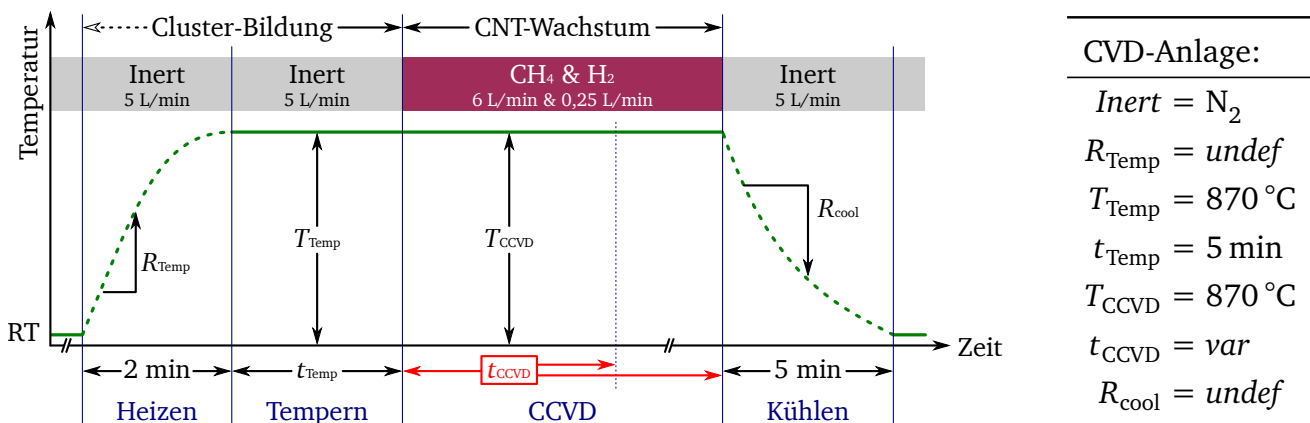


Abb. 49: Hinsichtlich der Dauer des CNT-Wachstums modifizierter Prozess unter Verwendung der CVD-Anlage, wobei für t_{CCVD} Werte von 10 min und 20 min gewählt wurden.

2,4 μm aufweist. Nach einer CNT-Wachstumsdauer von 20 min wird eine höhere Anzahl von 32 CNTs mit einer durchschnittliche Länge von ebenfalls 0,5 μm gewachsen; auch hier kann eine überdurchschnittlich lange CNT beobachtet werden, die in diesem Fall eine Länge von 2,1 μm hat. Die Messungen zeigen damit für die längere Wachstumsdauer eine deutlich größere Menge an CNTs, wobei die verdoppelte Dauer gut mit einer verdoppelten Anzahl an CNTs korreliert. Hinsichtlich der CNT-Länge kann keine Korrelation festgestellt werden, da für beide Wachstumsdauern die durchschnittliche Länge jeweils gleich ist und in beiden Fällen jeweils eine überdurchschnittlich lange CNT auftritt.

Aufgrund der Beobachtungen in Abb. 50 kann von den oben beschriebenen Möglichkeiten zunächst Grenzfall 3 ausgeschlossen werden, da — zumindest für die hier gewählten Wachstumsdauern — keine Sättigung des CNT-Wachstums feststellbar ist. Auch Grenzfall 1 kann ausgeschlossen werden, da die durchschnittliche CNT-Länge für beide Wachstumsdauern gleich ist und sich auch die beiden jeweils überdurchschnittlichen langen CNTs kaum unterscheiden. Folglich beschreibt Grenzfall 2 das CNT-Wachstum am zutreffendsten, der besagt, dass nach Erreichen einer gewissen CNT-Länge das Wachstum einer neuen CNT gegenüber der weiteren Verlängerung einer bereits vorhandenen CNT energetisch bevorzugt wird.

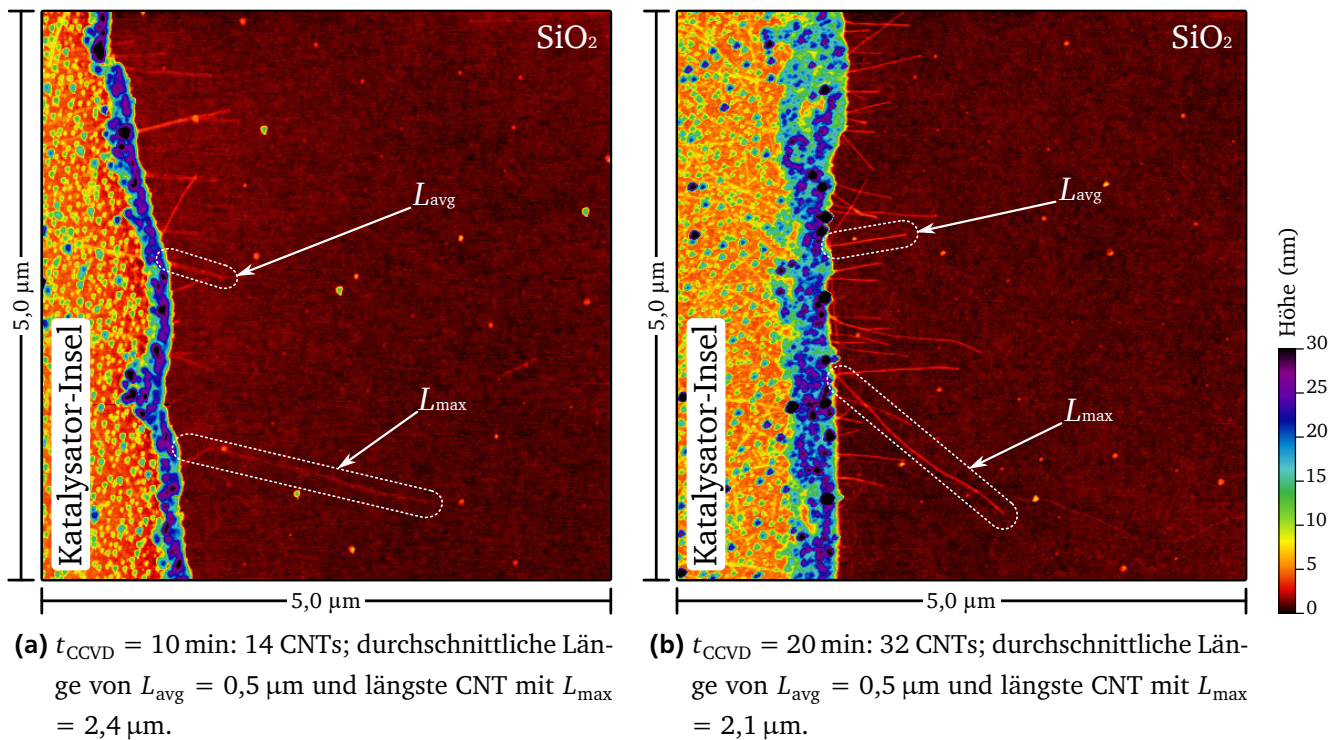


Abb. 50: Im Randbereich der Katalysator-Inseln durchgeführte AFM-Messungen auf Proben, die mit dem hinsichtlich t_{CCVD} modifizierten Prozess gemäß Abb. 49 hergestellt wurden. Die Messungen zeigen, dass eine längere Wachstumsdauer in einer höheren Anzahl an CNTs resultiert, während die durchschnittliche Länge nicht beeinflusst wird. (Der „Überstand“ bzw. das *Fencing* am Rand der Katalysator-Insel ist ein Artefakt der *Lift off*-Strukturierung.)

7.5 Einfluss der Cluster-Bildungs-Temperatur

Als nächster Parameter der Prozessführung soll die Temperatur der Cluster-Bildung T_{Temp} betrachtet werden. Der dafür genutzte Prozess ist in Abb. 51 skizziert, wobei Werte von 870, 920, 970 und 1020 °C für T_{Temp} gewählt wurden und der längere Wert von 20 min für die Dauer des CNT-Wachstums t_{CCVD} aus dem letzten Unterkapitel übernommen wird. Zu beachten ist hier, dass die Temperatur des CNT-Wachstums T_{CCVD} nicht geändert werden soll, d. h. dass der Wert von 870 °C des Basisprozesses beibehalten wird. Für die drei Werte von T_{Temp} , die oberhalb von T_{CCVD} liegen, wird damit ein kurzer Kühlschritt nach der Temperung notwendig. Die Dauer dieses Kühlschritts liegt im Bereich von 10 bis 20 s und ist von T_{Temp} abhängig, da anlagentechnisch bedingt ein freies Abkühlen stattfindet. Aufgrund der hohen verfügbaren Heizleistung der Anlage wird das Aufheizen zu Beginn des Prozesses von den unterschiedlichen Werten für T_{Temp} nicht beeinflusst und dauert nach wie vor rund 2 min.

Den Einfluss von T_{Temp} auf die Cluster-Bildung und das nachfolgende Wachstum der CNTs zeigen die AFM-Messungen in Abb. 52: Bei einer Temperung bei 870 °C bildet sich eine großen Menge an Clustern und CNTs, wobei die Cluster einen Durchmesser von etwa 30 nm haben. Die gewachsenen CNTs sind mit 3 nm recht dick und haben eine durchschnittliche Länge von 1 bis 1,5 μm . Eine Temperung bei 920 °C erzielt Cluster und CNTs, die mit 10 nm kleiner und mit 2 nm auch dünner sind. Außerdem sind diese CNTs mit 2 bis 3 μm länger. Durch eine Temperung bei 970 °C wird die Anzahl der Cluster und CNTs dramatisch reduziert. Die Abmessungen bleiben davon weitestgehend unbeeinflusst: sowohl die Größe der Cluster als auch der Durchmesser der CNTs liegen mit 15 nm bzw. 2 nm in der gleichen Größenordnung wie für 920 °C. Allerdings hat die Mehrheit der CNTs eine Länge von weniger als 1 μm . Nach einer Temperung bei 1020 °C treten plötzlich große Artefakte mit einer Fläche von etwa 200 nm \times 100 nm und einer Höhe

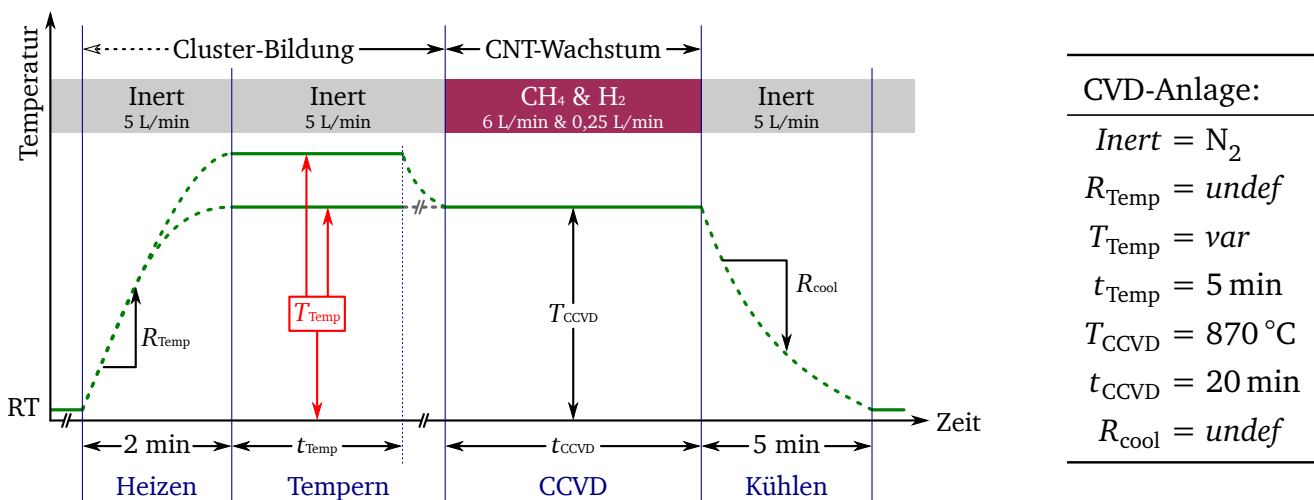


Abb. 51: Hinsichtlich der Temperatur der Cluster-Bildung modifizierter Prozess unter Verwendung der CVD-Anlage, wobei für T_{Temp} Werte von 870, 920, 970 und 1020 °C gewählt wurden.

von etwa 50 nm auf, deren Ursache hier nicht geklärt werden kann. Davon abgesehen ist das Ergebnis hinsichtlich der Cluster und CNTs mit dem für 970 °C vergleichbar.

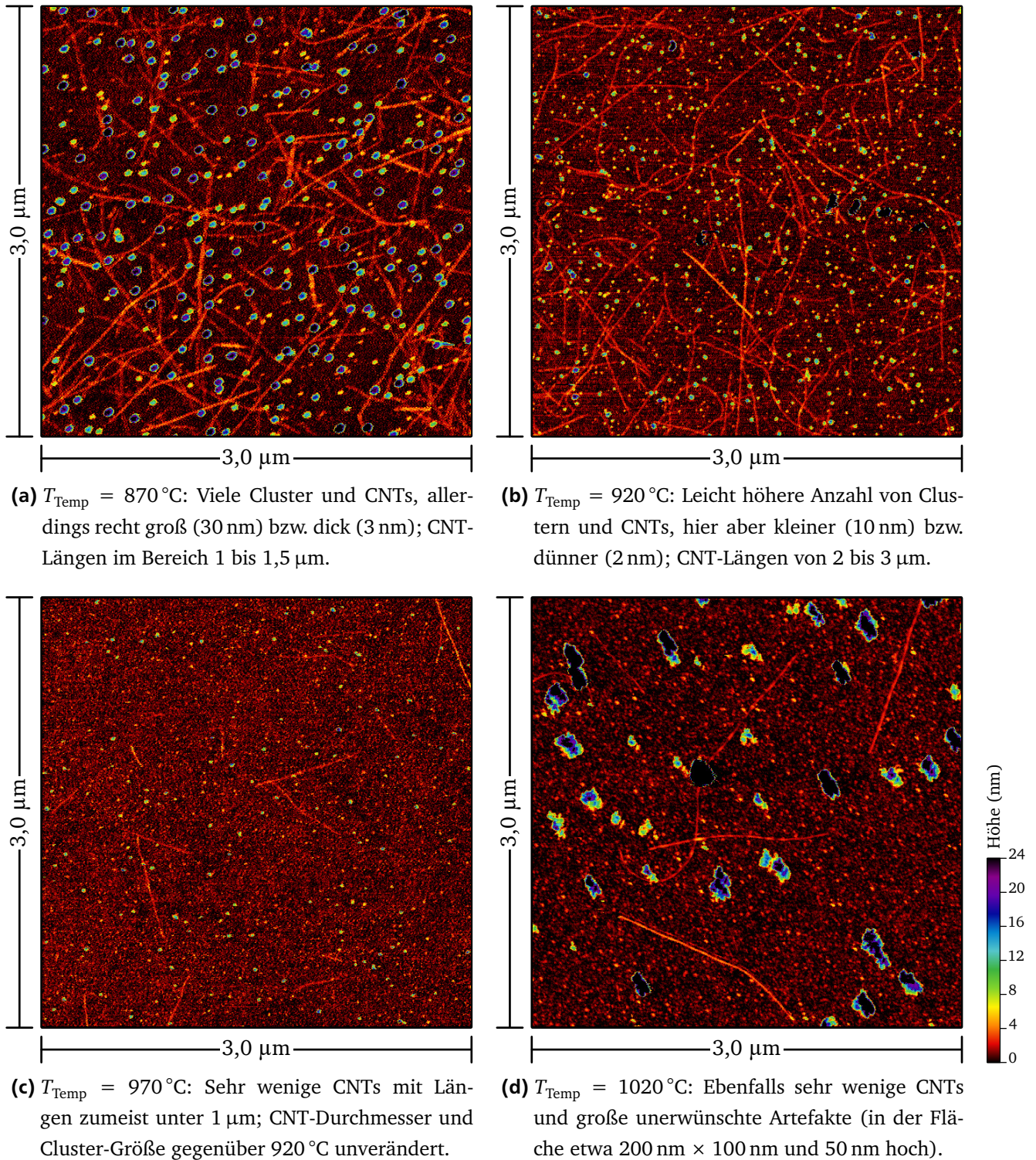


Abb. 52: Auf dem Katalysator durchgeführte AFM-Messungen auf Proben, die mit dem hinsichtlich T_{Temp} modifizierten Prozess gemäß Abb. 51 hergestellt wurden. Das vielversprechendste Ergebnis, d. h. viele und dünne CNTs, wird auf diesen AFM-Messungen für eine Cluster-Bildungs-Temperatur von 920 °C erzielt.

Was die eingangs geäußerte Zielsetzung angeht, möglichst viele und möglichst dünne CNTs zu wachsen, ist für die Cluster-Bildungs-Temperatur T_{Temp} der Wert von 920 °C, derjenige Wert mit dem besten Ergebnis. Für höhere Temperaturen werden nur sehr wenige und meist auch kurze CNTs gewachsen, während für niedrigere Temperaturen eine Zunahme des CNT-Durchmessers zu beobachten ist. Der Vergleich der AFM-Messungen in Abb. 52a und Abb. 52b bestätigt dabei nochmals die Theorie über den CNT-Wachstumsmechanismus, der einen Zusammenhang zwischen Cluster-Größe und CNT-Durchmesser unterstellt: Die bei 870 °C gebildeten Cluster haben eine Größe von 30 nm, während die gewachsenen CNTs einen Durchmesser von 3 nm aufweisen; bei der höheren Cluster-Bildungs-Temperatur von 920 °C werden sowohl kleinere Cluster gebildet als auch dünnere CNTs gewachsen, die eine Größe von 10 nm bzw. einen Durchmesser von 2 nm haben.

7.6 Einfluss der Heizrate der Cluster-Bildung

In diesem Unterkapitel soll schließlich der Einfluss der Heizrate R_{Temp} der Temperung auf die Cluster-Bildung untersucht werden. Hierzu werden zwei Versuche durchgeführt: Zunächst wird unter Verwendung der CVD-Anlage, die auch für die bisherigen Versuche genutzt wurde, der Einfluss zweier verschiedener Inertgase auf die Cluster-Bildung untersucht, wobei die jeweiligen Inertgase die Heizrate indirekt bestimmen. Für den zweiten Versuch wird die RTP-Anlage *SHS 1000G* des Herstellers *AST* genutzt, mit deren Hilfe eine Prozessierung möglich ist, die ein klar definiertes Temperatur-Profil aufweist.

Da in der RTP-Anlage aus technischen Gründen der CCVD-Schritt nicht durchgeführt werden kann, wird auch im ersten Versuch unter Verwendung der CVD-Anlage auf diesen verzichtet, um die Ergebnisse vergleichen zu können; die Prozessierung beschränkt sich hier also auf die Cluster-Bildung und spart das CNT-Wachstum aus. Zunächst soll dargelegt werden, warum für die CVD-Anlage eine Variation des Inertgases mit einer Variation der Heizrate der Temperung gleichgesetzt werden kann.

Zusammenhang zwischen Inertgas und Heizrate bei Verwendung der CVD-Anlage

Die hier genutzte CVD-Anlage hat einen Kaltwand-Reaktor, d. h. dass nur der Teller bzw. Substrathalter geheizt wird und damit das Gasvolumen über den Proben als Wärmesenke betrachtet werden muss. Für eine gegebene Heizleistung des Tellers bedeutet dies, dass die während der Temperung auf den Proben effektiv wirksame Heizrate niedriger ist. Um wie viel niedriger die tatsächliche Heizrate ist, wird durch die thermischen Eigenschaften des verwendeten Inertgases bestimmt. Für die zwei hier verwendeten Gase, Stickstoff und Argon, sind die Werte ihrer spezifischen Wärmekapazität und ihrer Wärmeleitfähigkeit in Tab. 3 zusammengestellt.

Tab. 3: Thermische Materialdaten für Stickstoff und Argon bei Raumtemperatur [79].

	spezifische Wärmekapazität c	Wärmeleitfähigkeit λ
Stickstoff	1,040 kJ/(kg·K)	25,83 mW/(m·K)
Argon	0,523 kJ/(kg·K)	17,72 mW/(m·K)

Dabei zeigt sich, dass sowohl die Wärmekapazität als auch die Wärmeleitfähigkeit von Stickstoff größer sind als die von Argon. Bei jeweils gleicher Heizleistung des Tellers muss die effektiv wirksame Heizrate auf den Proben in einer Argon-Atmosphäre also höher sein als in einer Stickstoff-Atmosphäre, da Argon vorhandene Wärme schlechter bzw. langsamer ableitet und sich selbst auch schneller aufheizt. Verglichen mit Stickstoff stellt Argon also während des Heizens die kleinere Wärmesenke dar, womit eine höhere Heizrate auf den Proben dadurch erzielt wird, dass Wärme langsamer an das Gasvolumen über den Proben abgegeben wird.

Diesen zunächst nur qualitativen Eindruck bestätigen die beiden berechneten Temperatur-Verläufe, die in Abb. 53 dargestellt sind, auch quantitativ. Die Berechnung erfolgt computer-unterstützt unter Verwendung der thermischen Materialdaten in Tab. 3 und der Wärmeleitungsgleichung gemäß Gl. 44; dabei sind ρ die Massedichte, c die spezifische Wärmekapazität und λ die Wärmeleitfähigkeit des jeweiligen Inertgases; $\partial T / \partial t$ die zeitliche Änderung der Temperatur; ∇ der Nabla-Operator; und Q die Heizleistung des Tellers [80, Kapitel 13].

$$\rho c + \frac{\partial T}{\partial t} + \nabla \cdot (-\lambda \nabla T) = Q \quad (44)$$

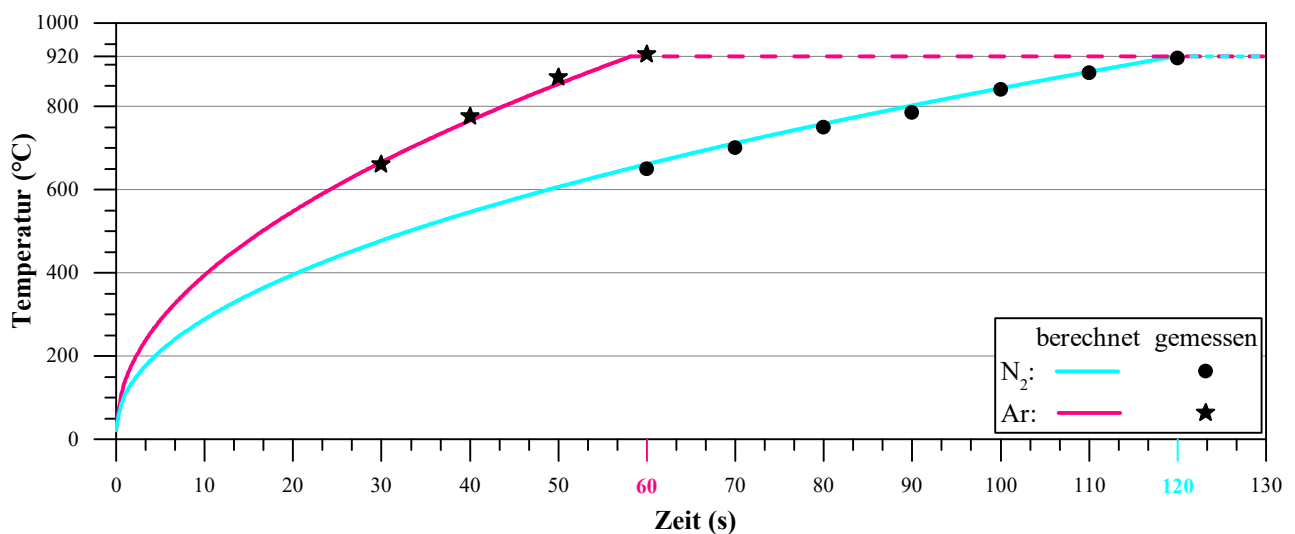


Abb. 53: Berechnete und gemessene Temperatur-Verläufe beim Heizen auf der Probenoberfläche in verschiedenen Inertgasen: Bei gleicher Heizleistung des Tellers der CVD-Anlage wird die Zieltemperatur der Temperung von $T_{\text{Temp}} = 920^\circ\text{C}$ in Stickstoff und Argon unterschiedlich schnell erreicht.

Abb. 53 zeigt klar, dass die Zieltemperatur der Temperung von $T_{\text{Temp}} = 920^\circ\text{C}$ in einer Argon-Atmosphäre bereits nach knapp 1 min erreicht ist, während es in Stickstoff mit 2 min doppelt so lang dauert. Die schwarzen Markierungen in Abb. 53 stellen in zwei Vorversuchen gemessene Temperaturen dar, wobei nur Werte oberhalb von 650°C angegeben sind, da die Temperaturerfassung der CVD-Anlage niedrigere Temperaturen nicht detektieren kann. Die gute Übereinstimmung der gemessenen Werte mit dem berechneten Kurven zeigt, dass die Berechnungen die realen Temperatur-Verläufe tatsächlich zutreffend abbilden.

Indirekte Beeinflussung der Heizrate in der CVD-Anlage durch das verwendete Inertgas

Im Gegensatz zu vielen anderen Forschungsgruppen, die oft Argon als Inertgas verwenden, wurden in dieser Arbeit bisher stets mit Stickstoff gearbeitet. Dies geschah aus anlagentechnischen Gründen und um zunächst die Vergleichbarkeit mit der Vorgängerarbeit zu wahren. Nach einigen Umbauten an der CVD-Anlage kann an dieser nun auch Argon anstelle von Stickstoff verwendet werden.

Der für die Variation des Inertgases genutzte Prozess ist in Abb. 54 skizziert. Auf Basis der im letzten Unterkapitel gefundenen Ergebnisse wird für die Cluster-Bildungs-Temperatur T_{Temp} der Wert von 920°C übernommen. Wie bereits erwähnt, umfasst dieser Prozess lediglich die Cluster-Bildung, da eine Vergleichbarkeit mit den noch folgenden Versuchen in der RTP-Anlage gegeben sein soll. Im Schritt „Heizen“ wird dabei über die unterschiedlichen thermischen Eigenschaften der genutzten Inertgase eine unterschiedliche Heizrate R_{Temp} erreicht, die aufgrund der Anlagenkonstruktion aber nach wie vor undefiniert ist. Die Berechnungen und gemessenen Temperaturen in Abb. 53 zeigen aber, dass die Cluster-Bildungs-Temperatur in Argon gegenüber Stickstoff nach kürzerer Zeit erreicht ist.

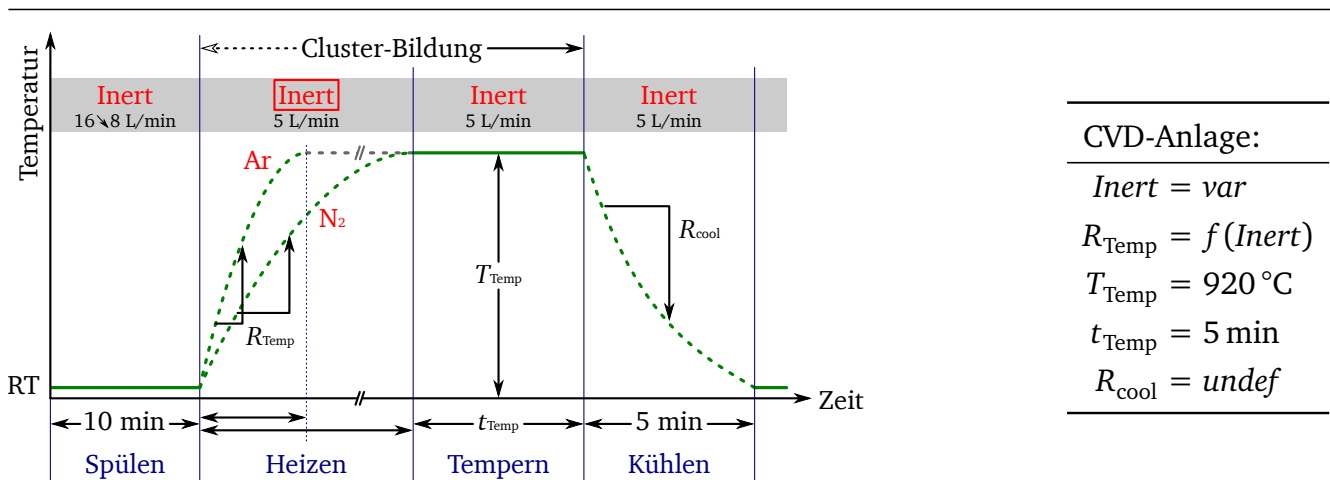


Abb. 54: Hinsichtlich des Inertgases modifizierter Prozess unter Verwendung der CVD-Anlage, wobei für $Inert$ Stickstoff und Argon genutzt wurden. Der CCVD-Schritt und damit das CNT-Wachstum entfallen hier, damit eine Vergleichbarkeit mit den noch folgenden Versuchen in der RTP-Anlage gegeben ist.

Abb. 55 zeigt anhand zweier AFM-Messungen den Einfluss des verwendeten Inertgases auf die Cluster-Bildung. Sofort wird augenscheinlich, dass bei einer Temperung in Argon eine deutlich höhere Anzahl an Clustern gebildet wird; auf der Probe mit Argon-Temperung finden sich rund 80 Cluster/ μm^2 , während es auf der Probe mit Stickstoff-Temperung lediglich 25 Cluster/ μm^2 sind. Anhand der beiden Profile, die aus den Messungen extrahiert wurden, zeigt sich, dass die Cluster bei einer Temperung in Stickstoff einen Durchmesser von üblicherweise 10 bis 16 nm erreichen. Kleinere Durchmesser im Bereich von 6 nm bis selten mehr als 10 nm werden bei einer Temperung in Argon erzielt. Hinsichtlich einer Prozessführung, mit der kleine Cluster gebildet werden, ist also eine Temperung in einem Inertgas vorteilhaft, das eine geringe spezifische Wärmekapazität und eine niedrige Wärmeleitfähigkeit aufweist. Dies ist hier für Argon der Fall, das gegenüber Stickstoff die kleinere Wärmesenke zur Probenoberfläche darstellt und somit eine höhere Heizrate bei der Cluster-Bildung zur Folge hat.

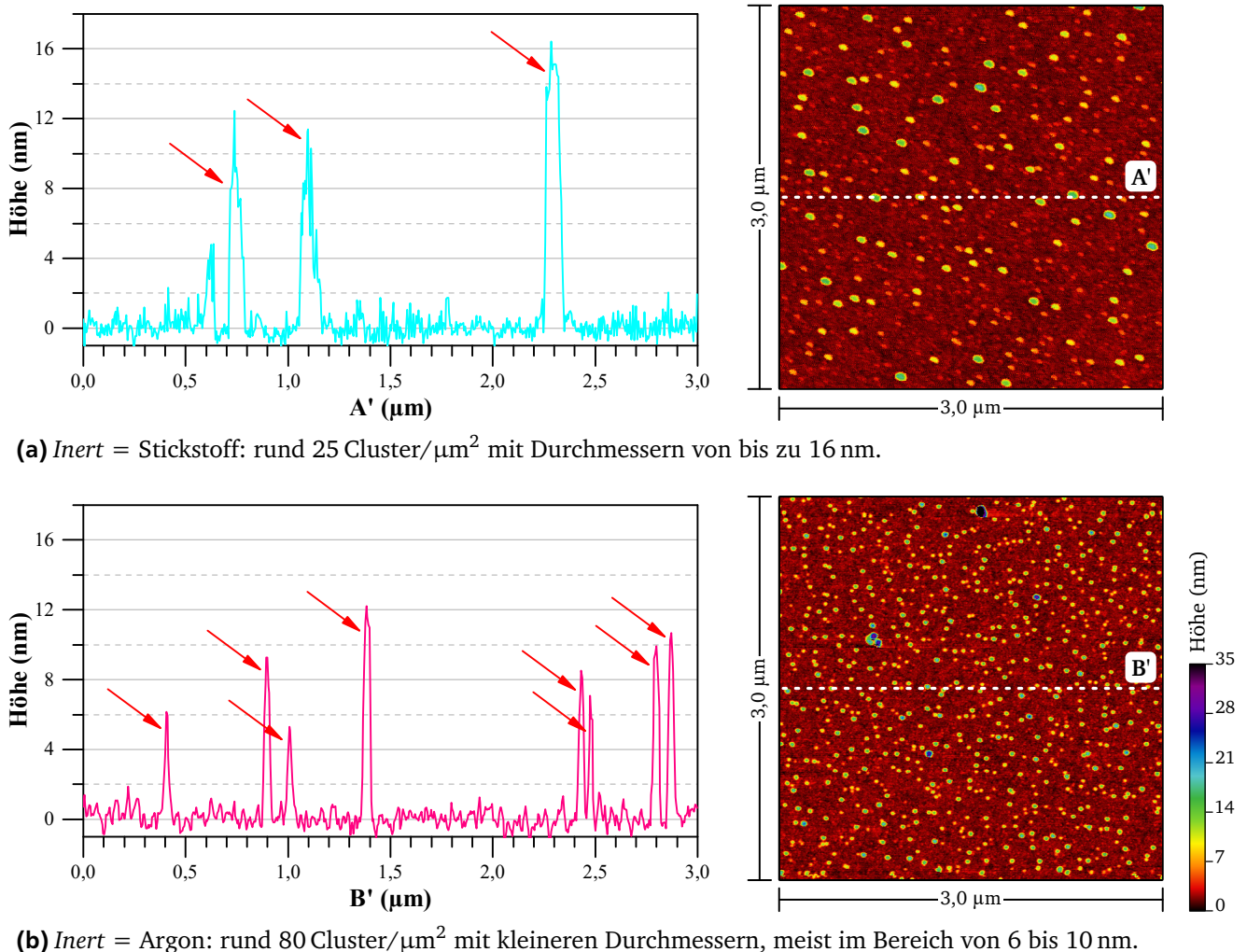


Abb. 55: Auf dem Katalysator durchgeführte AFM-Messungen (mit jeweils einem Profil in der Mitte der vermessenen Fläche) auf Proben, die mit dem hinsichtlich des Inertgases modifizierten Prozess gemäß Abb. 54 hergestellt wurden. Wie die Messungen zeigen, werden bei der Verwendung von Argon als Inertgas sowohl mehr als auch kleiner Cluster gebildet.

Direkte Beeinflussung der Heizrate durch Verwendung der RTP-Anlage

Aufgrund der Erkenntnis, dass durch die Verwendung von Argon als Inertgas indirekt eine höhere Heizrate der Cluster-Bildung erreicht wird und dass diese die Bildung kleiner Cluster begünstigt, soll hier nun der Einfluss der Heizrate R_{Temp} genauer untersucht werden. Dafür werden eine RTP-Anlage und der in Abb. 56 skizzierte Prozess genutzt, wobei die Heizrate direkt festgelegt werden kann. Die Prozessführung ist dabei den bisherigen Prozessen in der CVD-Anlage nachempfunden, um vergleichbare Ergebnisse zu erhalten. Den Vorteil der RTP-Anlage stellt die Möglichkeit einer präzisen Temperatur-Führung dar — besonders während des Schrittes „Heizen“; der Nachteil ist, dass in dieser Anlage der CCVD-Schritt nicht durchgeführt werden kann und deswegen das CNT-Wachstum entfallen muss.

Der in Abb. 56 dargestellte Prozess beginnt mit einem Spülschritt mit 10 min Dauer in Argon. Nach dem Spülen findet ein Vorheizen auf 500 °C statt, das 10 s dauert und der Stabilisierung der Temperatur dient. Ausgehend von dieser Temperatur wird dann mit konstanten Heizraten R_{Temp} im Bereich von 1,4 bis 28 K/s auf die Cluster-Bildungs-Temperatur T_{Temp} von 920 °C aufgeheizt, die nach Erreichen für $t_{\text{Temp}} = 5$ min gehalten wird. Danach wird mit einer definierten Rate R_{cool} von 2,4 K/s abgekühlt. Nach Unterschreiten von 350 °C erfolgt ein freies Abkühlen, da von der RTP-Anlage nur Temperaturen oberhalb von 350 °C detektiert werden können.

Abb. 57 zeigt AFM-Messungen für drei exemplarische Heizraten R_{Temp} von 3,5, 7 und 14 K/s samt jeweils einem Profil aus der Mitte der Messung. Bereits diese drei Messungen lassen einen deutlichen Trend erkennen: Wird die Heizrate erhöht, also das Aufheizen auf die Zieltemperatur in kürzer Zeit durchgeführt, nimmt die Anzahl an großen und mittleren Clustern mit einer Größe über 2 nm rapide ab. Dieser zunächst nur qualitative Eindruck wird auch quantitativ durch

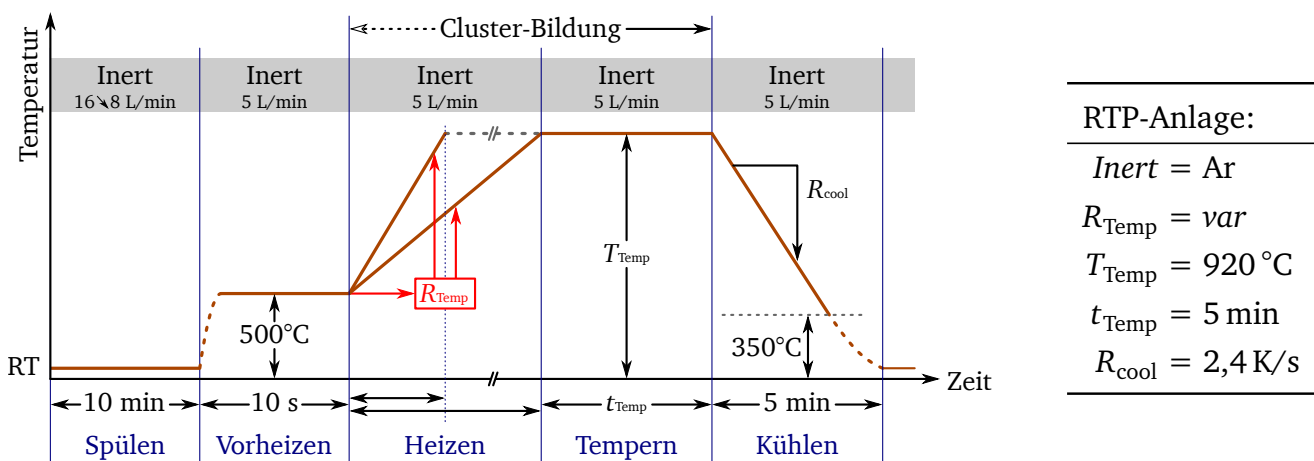


Abb. 56: Hinsichtlich der Heizrate der Cluster-Bildung modifizierter Prozess unter Verwendung der RTP-Anlage, wobei für R_{Temp} Werte im Bereich von 1,4 bis 28 K/s gewählt wurden. Zu beachten ist, dass die RTP-Anlage nur Temperaturen oberhalb von 350 °C detektieren kann, sodass beim Kühlen nach Unterschreitung dieser Temperatur wie in der CVD-Anlage ein freies Abkühlen stattfindet. Die Kühlrate R_{cool} ist so gewählt, dass das gesamte Kühlen wie bei den bisherigen Prozessen rund 5 min dauert.

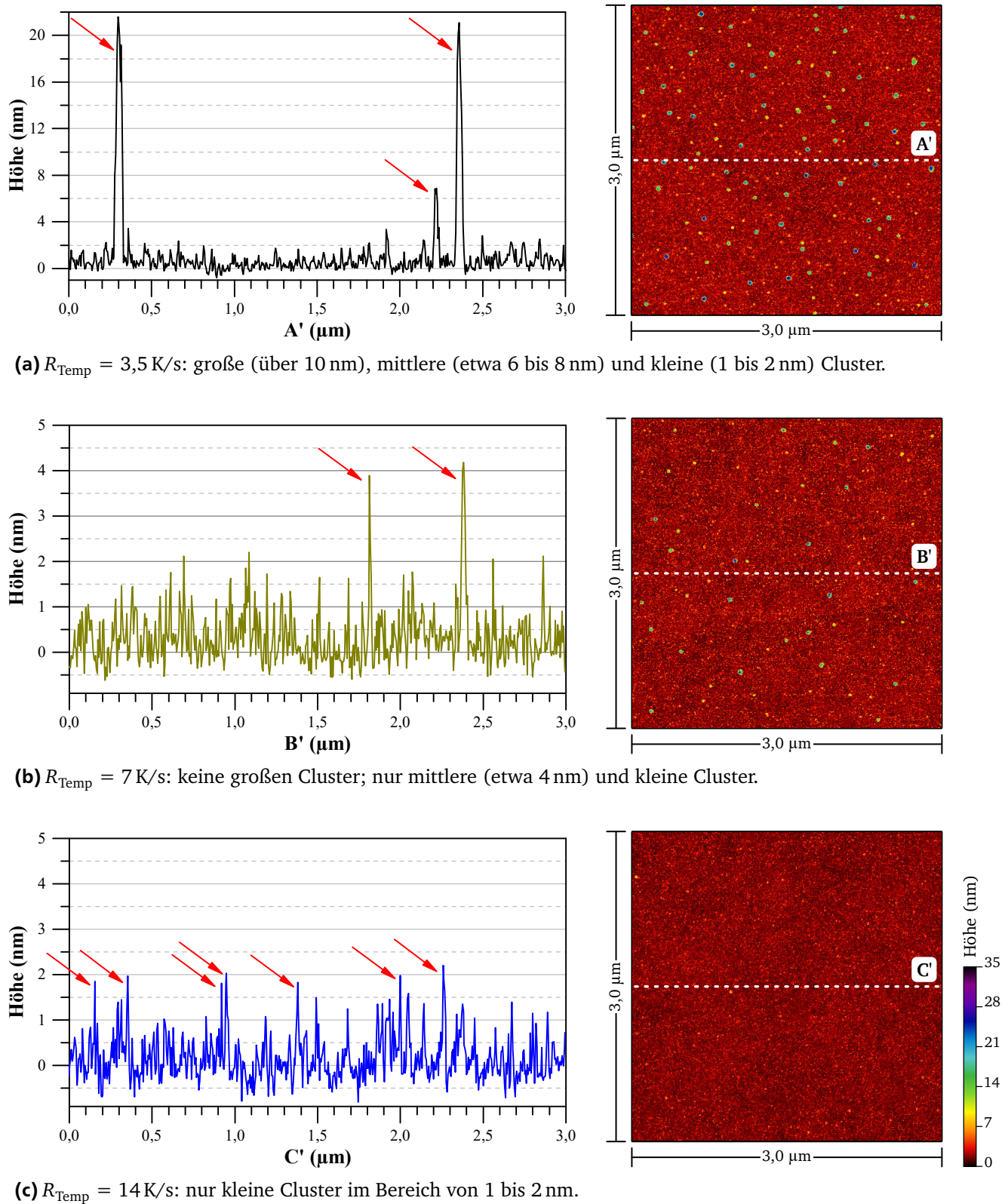


Abb. 57: Auf dem Katalysator durchgeführte AFM-Messungen (mit jeweils einem Profil in der Mitte der vermessenen Fläche) auf Proben, die mit dem hinsichtlich R_{Temp} modifizierten Prozess in der RTP-Anlage gemäß Abb. 56 hergestellt wurden. Die Messungen zeigen, dass für höhere Heizraten kleinere Cluster gebildet werden. Weitere Heizraten und detaillierte Angaben zu Cluster-Größen und -Anzahl siehe Tab. 4.

Tab. 4: Cluster-Größen und -Anzahl für verschiedene Heizraten R_{Temp} der Cluster-Bildung, ausgewertet auf Basis von zehn Profilen pro Heizrate bzw. AFM-Messung.

R_{Temp} [K/s]	Cluster-Größe [nm]					\emptyset
	≤ 2	≤ 5	≤ 10	≤ 15	≤ 20	
	„klein“	\Leftarrow „mittel“ \Rightarrow	\Leftarrow „groß“ \Rightarrow			
1,4	176	19	4	7	–	2,87
1,75	185	39	7	1	1	2,88
2,3	162	30	9	3	–	2,99
3,5	226	40	8	1	1	2,78
4,6	209	6	–	–	–	2,08
7	252	27	2	–	–	2,35
14	230	10	–	–	–	2,13
28	179	2	–	–	–	2,03

Tab. 4 bestätigt. Die Bestimmung der Anzahl und Größe der Cluster erfolgt dabei auf Basis der Auswertung von jeweils zehn Profilen pro untersuchter Heizrate bzw. AFM-Messung. Besonders im Bereich der großen Cluster ist deutlich zu erkennen, dass deren Auftreten mit steigender Heizrate zurückgeht. Für Heizraten größer als 3,5 K/s sind praktisch nur noch kleine und eine überschaubare Anzahl an mittleren Clustern nachweisbar. Dies bestätigt zusätzlich die als Indikator berechnete Cluster-Größe \emptyset , die für hohe Heizraten kleinere Werte als für geringere Heizraten annimmt. Dass die konkreten Werte in Tab. 4 den beschriebenen Zusammenhang nicht vollständig abbilden, ist der Art der Probennahme und der Auswertung geschuldet: Für die Berechnung des Indikators wurde die jeweilige Obergrenze der Größenintervalle herangezogen, was zu einer Überschätzung der Cluster-Größe führt. Dies wird dadurch verschärft, dass große und mittlere Cluster eine höhere Nachweiswahrscheinlichkeit besitzen, da diese auf dicht benachbarten Profilen der AFM-Messung immer wieder abgebildet werden, während kleine Cluster nur auf einem der Profile sichtbar werden. Da aufgrund des nötigen Aufwandes nur zehn Profile pro Messung ausgewertet wurden, erscheint der Anteil kleiner Cluster somit reduziert. Auch weist die Oberfläche des Katalysators nach der Cluster-Bildung eine mittlere Rauheit von rund 0,5 nm auf, weswegen kleine Cluster potentiell nicht vom Untergrund unterschieden werden können. Auch dies führt zu einer prominenteren Darstellung großer Cluster.

Dass höhere Heizraten, ob nun direkt mit Hilfe der RTP-Anlage eingestellt oder indirekt über das Inertgas erzeugt, kleinere Cluster zur Folge haben, kann mit einem der OSTWALD-Reifung ähnlichen Mechanismus erklärt werden. Die OSTWALD-Reifung ist ein kolloidchemischer Prozess, demzufolge sich in einer Gruppe von Kolloiden verschiedener Größe die Anzahl der Kolloiden

mit der Zeit verringert, indem ein Materialtransport von kleinen zu größeren Kolloiden stattfindet, bis sich die kleinen Kolloiden schließlich auflösen [81]. Verursacht wird dieser Materialtransport durch den höheren Dampfdruck, den kleinere Kolloiden durch ihre stärker gekrümmte Oberfläche besitzen. Dass die OSTWALD-Reifung auch für polykristalline Metalle und Legierungen, und damit für Feststoffe, Relevanz besitzt, wurde unter anderen von HILLERT und ORIANI beschrieben [82, 83]. Die OSTWALD-Reifung ist ein von selbst ablaufender Prozess, d. h. er benötigt keine Aktivierungsenergie und läuft damit prinzipiell auch bei niedrigen Temperaturen ab. Da der Dampfdruck aber temperaturabhängig ist, lässt sich leicht folgern, dass sich die OSTWALD-Reifung durch erhöhte Temperaturen beschleunigen lässt.

Angewandt auf die Cluster-Bildung bedeutet die OSTWALD-Reifung, dass die hier angestrebten kleinen Nickel-Cluster einen instabilen Zustand darstellen, der sich mit der Zeit und besonders bei erhöhter Temperatur in Richtung eines stabileren, aber unerwünschten, Zustands mit großen Clustern verschieben wird. Soll also die Bildung großer Cluster vermieden werden, müssen hohe Heizraten gewählt werden, um die Dauer mit hohen Temperaturen möglichst kurz zu halten.

7.7 Zusammenfassung der Ergebnisse und optimierte Prozessführung

Beginnend mit der Untersuchung des Einflusses der Sauerstoff-Plasmabehandlung des Katalysators auf die Cluster-Bildung und auf das CNT-Wachstum wurden in diesem Kapitel verschiedene Parameter der Prozessführung mittels topologischer Messungen untersucht. Als eigentliche Parameter der Prozessführung während der Temperung und des CCVD-Schritts wurden ausgehend vom Basisprozess die CNT-Wachstumsdauer t_{CCVD} , die Cluster-Bildungs-Temperatur T_{Temp} und die Heizrate der Cluster-Bildung R_{Temp} betrachtet. Hierbei wurden unter der Zielsetzung einen Prozess zu finden, mit dessen Hilfe eine hohe Anzahl möglichst langer und möglichst dünner CNTs *in situ* gewachsen wird, der jeweils beste Wert für den gerade betrachteten Parameter für die nächsten Unterkapitel bzw. Untersuchungen übernommen, und so der Basisprozess sukzessive optimiert. Auf Basis der in den letzten vier Unterkapiteln durchgeführten AFM-Messungen konnten folgende optimierte Parameter der Prozessführung gefunden werden:

- **Sauerstoff-Plasmabehandlung des Katalysators:**

Die Sauerstoff-Plasmabehandlung beeinflusst die Cluster-Bildung und das CNT-Wachstum positiv. Auf Plasma-behandelten Proben werden mehr und gleichzeitig dünnere CNTs gewachsen; die CNT-Durchmesser liegen im Bereich von 1 bis 2 nm.

- **CNT-Wachstumsdauer:**

Für die längere Wachstumsdauer t_{CCVD} von 20 min werden mehr CNTs gewachsen als für die kürzere Dauer. Die Erhöhung der Wachstumsdauer beeinflusst allerdings nicht die

durchschnittliche Länge von 0,5 μm der CNTs, wobei unabhängig von der Dauer wenige überdurchschnittlich lange CNTs mit einer Länge im Bereich von 2 bis 2,5 μm auftreten.

- **Cluster-Bildungs-Temperatur:**

Eine Cluster-Bildung bei $T_{\text{Temp}} = 920^\circ\text{C}$ führt hier zum besten Ergebnis mit CNT-Durchmessern von 2 nm und CNT-Längen im Bereich von 2 bis 3 μm . Bei höheren Temperaturen werden signifikant weniger CNTs gewachsen, die auch kürzer sind; hin zu niedrigeren Temperaturen nimmt der Durchmesser der CNTs merklich zu.

- **Inertgas während der Cluster-Bildung:**

Eine Cluster-Bildung in Argon als Inertgas führt gegenüber einer Temperung in Stickstoff zu kleineren und einer höheren Anzahl von Clustern mit Größen im Bereich von 6 bis 10 nm. Ausschlaggebend ist an dieser Stelle die höhere Heizrate der Temperung R_{Temp} , die durch die geringere spezifische Wärmekapazität und geringere Wärmeleitfähigkeit von Argon indirekt erzeugt wird.

Da in der RTP-Anlage der CCVD-Schritt nicht durchgeführt werden kann, können die Ergebnisse zur Untersuchung der direkt bestimmten Heizrate der Cluster-Bildung hier nicht weiter genutzt werden. Denkbar wäre zwar, für die Cluster-Bildung die RTP-Anlage zu nutzen, die Proben im Anschluss in die CVD-Anlage zu transferieren und dort das CNT-Wachstum durchzuführen. Dies würde aber ein weiteres undefiniertes Aufheizen in der CVD-Anlage bedeuten, was die zuvor mit einer definierten Heizrate R_{Temp} von zum Beispiel 28 K/s gebildeten Cluster wieder verändern dürfte.

Ausgehend vom Basisprozess gemäß Abb. 47 (Seite 70) wird also ein wie folgt optimierter Prozess im nächsten Kapitel Anwendung finden (vgl. auch Abb. 58): Die Sauerstoff-Plasmabehandlung des Katalysators wird durchgeführt; die Temperung zur Cluster-Bildung findet in der CVD-Anlage in Argon bei T_{Temp} bei 920°C für eine Dauer t_{Temp} 5 min statt; nach einem kurzen Zwischenkühlen folgt das CNT-Wachstum bei $T_{\text{CCVD}} = 870^\circ\text{C}$ für $t_{\text{CCVD}} = 20$ min.

8 CNTFETs mit Al/Ni-Katalysator

8.1 Prozess zur Herstellung der FETs

Auf Basis der im letzten Kapitel durchgeführten topologischen Untersuchungen und der optimierten Prozessführung hinsichtlich der Cluster-Bildung und des CNT-Wachstums wurde eine Reihe von Wafern prozessiert. Der verwendete Prozess nutzt den bereits vorgestellten Aluminium/Nickel-Katalysator mit den folgenden Eckdaten:

- hoch Bor-dotierte Silizium-Substrate
- 100 nm thermisch gewachsenes Trocken-Oxid
- strukturierter Katalysator bestehend aus 5 nm Aluminium und 1 nm Nickel
- Sauerstoff-Plasmabehandlung des Katalysators
- Temperung in Argon bei $T_{\text{Temp}} = 920^\circ\text{C}$ für $t_{\text{Temp}} = 5\text{ min}$
- CCVD mit Methan und Wasserstoff bei $T_{\text{CCVD}} = 870^\circ\text{C}$ für $t_{\text{CCVD}} = 20\text{ min}$ (vgl. Abb. 58)
- *Source/Drain*-Kontakte mit 10 nm Palladium und 100 nm Nickel-Capping.
- Unterseite der Substrate als *Gate*-Kontakt (rückseitiges Oxid selektiv entfernt)

Die Prozessführung während der Temperung und des CCVD-Schritts stellt dabei eine Modifikation des Basisprozesses der Vorgängerarbeit dar und ist in Abb. 58 skizziert. Hinsichtlich der verschiedenen Parameter der Prozessführung für die Cluster-Bildung und das CNT-Wachstum sei auf Kapitel 7.7 (Seite 83) verwiesen.

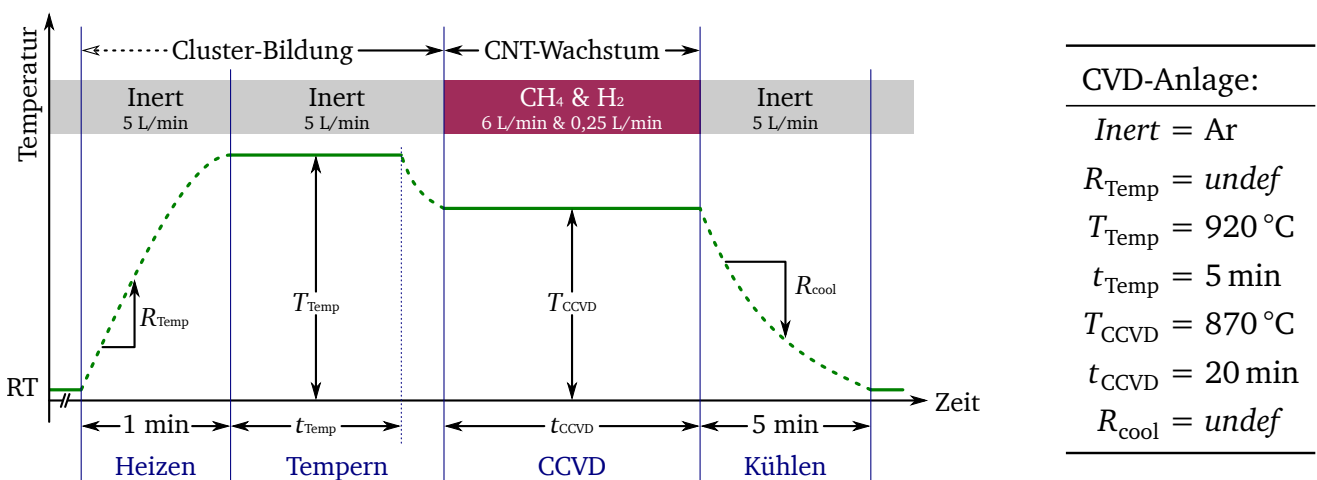


Abb. 58: Details der Prozessführung während der Temperung und des CCVD-Schritts für die Cluster-Bildung und das CNT-Wachstum zur Herstellung der CNTFETs mit Aluminium/Nickel-Katalysator.

8.2 Elektrische Untersuchung von ausgewählten FETs

Bevor die eigentliche Auswertung vorgestellt wird, muss hier festgehalten werden, dass bei den durchgeführten Messungen ein unerwartetes und schwerwiegendes Zuverlässigkeitsproblem aufgetreten ist. Bei nahezu allen Bauelementen kommt es bereits bei *Gate*-Spannungen V_{GS} von unterhalb 10 V zum elektrischen Durchbruch des Dielektrikums. Dies entspricht bei der hier verwendeten Oxiddicke von 100 nm einer Durchbruchfeldstärke von weniger als 1 MV/cm und liegt damit weit entfernt von den Werten von 10 MV/cm, die für ein thermisch gewachsenes Trocken-Oxid üblicherweise erwartet werden können. Die Schädigung des Dielektrikums ist zudem so umfassend, dass die Ausbeute an Bauelementen ohne dielektrischen Durchbruch bei $V_{GS} = \pm 10$ V bei gerade einmal 5 % pro Wafer liegt. Kapitel 9 ist deswegen der Suche nach der Ursache der *Gate*-Durchbrüche gewidmet. Im Folgenden müssen sich die Untersuchungen auf die wenigen Bauelemente stützen, die ohne dielektrischen Durchbruch vermessen werden konnten.

Die Bauelemente, die eine Vermessung zuließen, lassen sich hinsichtlich ihres elektrischen Verhaltens in zwei Gruppen einteilen: Entweder zeigen die Bauelemente ein gutes Schaltverhalten, d. h. ein hohes *On/Off*-Verhältnis und einen niedrigen Strom I_{off} im hochohmschen bzw. *Off*-Zustand, oder sie zeigen einen hohen Strom I_{on} im niederohmschen bzw. *On*-Zustand und ein geringes *On/Off*-Verhältnis. Beide Gruppen seien hier durch die Bauelemente FET 1 und FET 2 repräsentiert, die gemeinsam gemäß Abb. 58 prozessiert und unter den gleichen Bedingungen gemessen wurden. Sowohl FET 1 und als auch FET 2 sind FETs mit 32 Fingerpaaren und haben damit eine nominale Kanalweite $W_{Kanal|nom}$ von $32 \times 180 \mu\text{m} = 5760 \mu\text{m}$. FET 1 hat eine nominale Kanallänge $L_{Kanal|nom}$ von $5 \mu\text{m}$ und FET 2 von $2 \mu\text{m}$. Um eine Einordnung dieser ersten Ergebnisse vornehmen zu können, wurden vergleichend an einem repräsentativen Bauelement aus der Vorgängerarbeit [6] Messungen durchgeführt. Dieses Vergleichsbauelement FET R hat eine nominale Kanalweite und -länge von jeweils $5 \mu\text{m}$. Zu beachten ist dabei, dass FET R mit 40 nm gegenüber FET 1 und FET 2 auf einem dünneren Siliziumdioxid hergestellt wurde und damit eine bessere elektrostatische Kontrolle des *Gates* auf den Kanal aufweist.

Unterschwelkenennlinien

Die in Abb. 59a und Abb. 59c gezeigten Unterschwelkenennlinien lassen für FET 1 und FET 2 unipolares p-MOS-Verhalten erkennen, d. h. der *On*-Zustand stellt sich bei negativen *Gate*-Spannungen V_{GS} ein. Für die Messung der Kennlinien wurde V_{GS} bei einem konstantem V_{DS} von -500 mV beginnend bei -10 V auf $+10$ V erhöht. Die aus diesen Kennlinien ermittelten Größen sind auch in Tab. 5 zusammengefasst. FET 1 zeigt dabei einen *On*-Strom I_{on} von $1 \mu\text{A}$ und einen *Off*-Strom I_{off} von 50 pA und damit ein *On/Off*-Verhältnis von rund 4,5 Größenordnungen und einen *On*-Widerstand R_{on} von 500 k Ω . Die Unterschwelkensteigung S_{subthr} beträgt

rund 800 mV/dec. FET 2 dagegen zeigt einen deutlich höheren *On*-Strom von 80 μA , aber auch einen signifikant höheren *Off*-Strom von 30 μA , was einem *On/Off*-Verhältnis von lediglich einem Faktor 2,66 entspricht. Abb. 59d zeigt die Unterschwellenkennlinie, die an FET R gemessen wurde. Dieses Bauelement hat einen *On*-Strom von 600 nA und einen *Off*-Strom von 3 pA, was einem *On/Off*-Verhältnis von rund 5,5 Größenordnungen und einem *On*-Widerstand R_{on} von 666 k Ω entspricht. Die Unterschwellensteigung S_{subthr} liegt hier bei rund 300 mV/dec.

Sowohl FET 1 als auch FET R zeigen Werte für R_{on} , die einen Faktor 100 größer sind als das theoretische Minimum von 6,5 k Ω für SWNTs [84]. Ursache dafür können parasitäre Widerstände an den *Source/Drain*-Kontakten und strukturelle Defekte der CNTs sein. Parasitäre Widerstände an den *Source/Drain*-Kontakten, also am Übergang von der Palladium-Schicht zur CNT, können hier dadurch entstehen, dass während der Abscheidung des Palladiums eine Abschattung der CNT durch Nickel-Cluster vorliegt. Da das Palladium hier nicht getempert wird und generell schlechte Haftungseigenschaften gezeigt hat, verhindert eine solche Abschattung einen vollständigen Kontakt der Palladium-Schicht zur CNT und führt somit zu einem erhöhten elektrischen Widerstand. Wie in den Grundlagen zu Kohlenstoffnanoröhren in Kapitel 2.2 (Seite 5) dargelegt, basieren die elektronischen Eigenschaften von CNTs auf der hexagonalen periodischen Struktur von Graphen. Strukturelle Defekte innerhalb der CNT unterbrechen diese Periodizität und stören somit die Bewegung der Ladungsträger, sodass deren Mobilität reduziert wird, was zu einer Reduzierung des *On*-Stroms und damit zum einem höheren Widerstand des FET führt.

Im Grundlagen-Kapitel zu Feldeffekt-Transistoren wurde zur Bestimmung der Schwellspannung eines MOSFET auf Seite 23 nach SCHRODER die Methode der linearen Extrapolation genannt. Dieses Verfahren ist allerdings nur auf Messungen mit vernachlässigbaren Serienwiderständen anwendbar, wovon hier nicht ausgegangen werden kann. Außerdem basiert das Verfahren auf der Definition der Schwellspannung eines MOSFET, also der Bildung eines leitfähigen Inversionskanals zwischen *Source* und *Drain* — was für die hier betrachteten CNTFETs nicht zutreffend ist. Deswegen soll hier, wie in der Vorgängerarbeit, eine abweichende Methode mittels graphischer Bestimmung der Schwellspannung und der Unterschwellensteigung genutzt werden: In eine Unterschwellenkennlinie werden wie in Abb. 59a und Abb. 59d zwei Geraden eingezeichnet; die erste horizontal, auf Höhe des *On*-Stromes, die zweite schräg, als Tangente an den Verlauf der Kurve im Übergangsbereich von *On*- zu *Off*-Zustand; der V_{GS} -Wert des Schnittpunkts

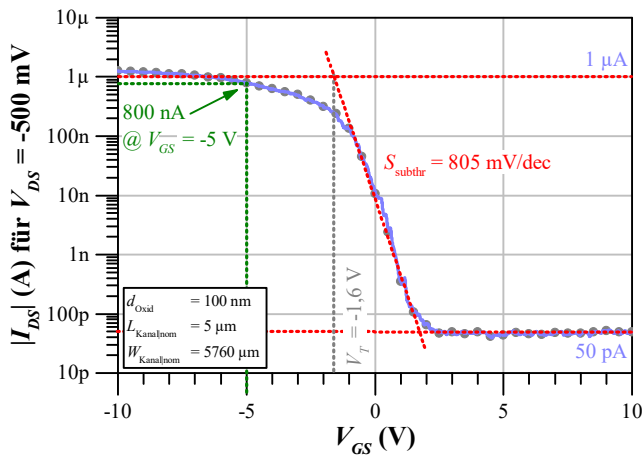
Tab. 5: Aus den Unterschwellenkennlinien in Abb. 59 für FET 1, FET 2 und FET R ermittelte Größen.

	d_{Oxid}	I_{on}	I_{off}	<i>On/Off</i>	R_{on}	S_{subthr}	V_T
FET 1	100 nm	1 μA	50 pA	2×10^4	500 k Ω	805 mV/dec	−1,6 V
FET 2	100 nm	80 μA	30 μA	2,66	6,25 k Ω	—	—
FET R	40 nm	600 nA	3 pA	2×10^5	666 k Ω	305 mV/dec	−1,5 V

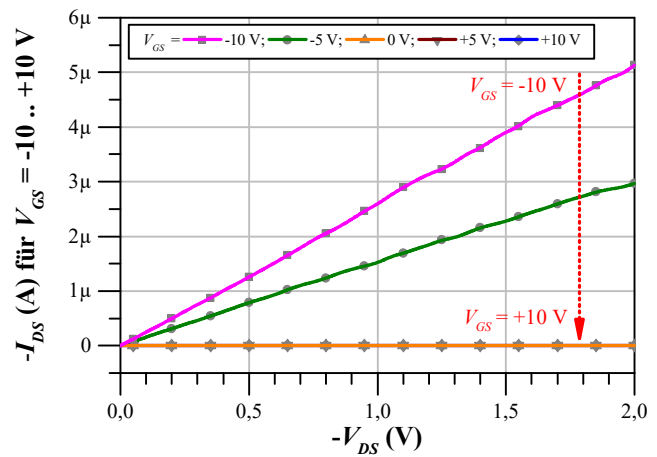
dieser beiden Geraden sei dann schließlich die Schwellspannung V_T des CNTFET. Die zweite Gerade dient dabei gleichzeitig der Bestimmung der Unterschwellensteigung S_{subthr} . Auf diese Weise können für FET 1 und FETR Schwellspannungen von $-1,6\text{ V}$ bzw. $-1,5\text{ V}$ bestimmt werden.

Ausgangskennlinien

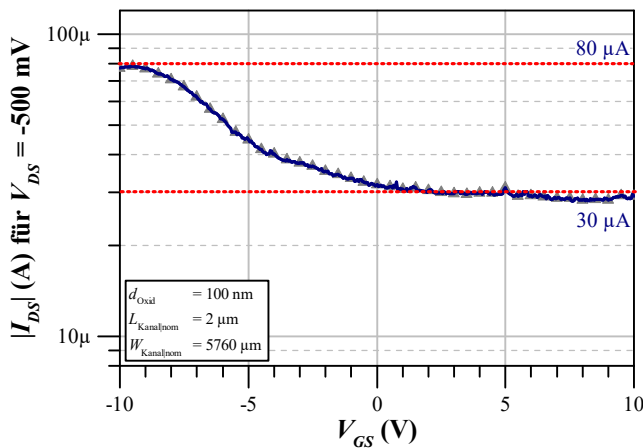
Für FET 1 wurde zusätzlich eine Ausgangskennlinienschar aufgenommen, die in Abb. 59b gezeigt ist. Die in Intervallen von 5 V der Gate-Spannung V_{GS} gemessenen Kurven zeigen dabei bei $V_{DS} = -500\text{ mV}$ eine gute Übereinstimmung in I_{DS} mit der Unterschwellenkennlinie. Erwartungsgemäß können hier nur zwei der eigentlich drei vom MOSFET bekannten Arbeitsbereiche beobachtet werden. Für $V_{GS} \geq 0\text{ V}$ befindet sich das Bauelement im Sperrbereich, d. h. es kann kein nennenswerter Source/Drain-Strom I_{DS} fließen. Für die beiden Kurven mit negativen Wer-



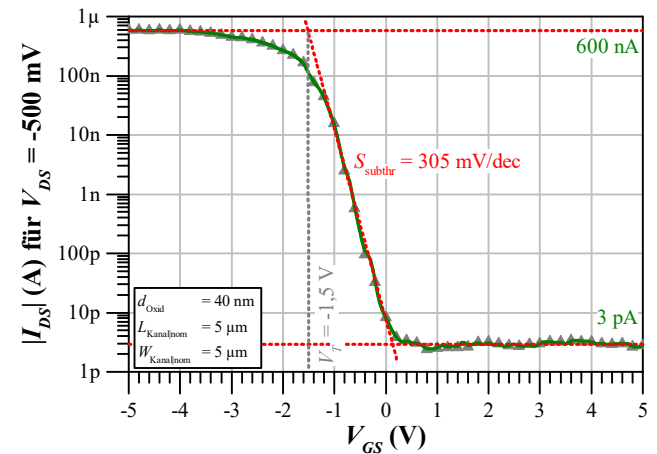
(a) FET 1: Unterschwellenkennlinie



(b) FET 1: Ausgangskennlinienschar



(c) FET 2: Unterschwellenkennlinie



(d) FETR: Unterschwellenkennlinie

Abb. 59: Elektrische Charakterisierung von FET 1 und FET 2: Aufgrund der linearen Darstellung überlappen sich in (b) die Kurven für $V_{GS} \geq 0\text{ V}$ teilweise. Zum Vergleich eine neue Messung, FETR in (d), auf einem alten Wafer aus [6].

ten für V_{GS} ist das Bauelement im linearen Bereich, d. h. es fließt ein von V_{DS} abhängiger Strom. Der hier nicht zu beobachtende Arbeitsbereich ist der Sättigungsbereich, der für $V_{DS} \geq V_{DS,sat}$ erreicht werden sollte. Im Sättigungsbereich hätte eine weitere Erhöhung von V_{DS} keine weitere Erhöhung des Stroms zur Folge.

Das Fehlen des Sättigungsbereichs kann in zweifacher Weise erklärt werden. Erstens wird die Sättigung des *Source/Drain*-Stroms beim MOSFET durch die Abschnürung (*pinch off*) des Inversionskanals verursacht. Das bedeutet, dass Ladungsträger auf dem Weg von *Source* nach *Drain* den abgeschnürten, nicht leitfähigen Bereich des Kanals durch Tunneln überwinden müssen. Dieser Tunnel-Strom ist für $V_{DS} \geq V_{DS,sat}$ konstant, sodass eine weitere Erhöhung des Stroms nicht erreicht werden kann. Da CNTFETs aber keinen Inversionskanal aufweisen, kommt es für sie zu keiner Abschnürung des Kanals und damit auch zu keiner Sättigung des *Source/Drain*-Stroms. Zweitens, selbst wenn FET 1 *per se* eine Sättigung zeigen würde, wäre diese hier wegen der hohen Kontakt-Widerstände nicht zu beobachten. Dies zeigen die Ergebnisse der Arbeit von HOENLEIN ET AL. [85], die demonstrieren, dass hohe Widerstände an den *Source/Drain*-Kontakten ein nicht-ideales Verhalten von CNTFETs zur Folge hat. Demnach wird eine Sättigung je nach Höhe des Kontakt-Widerstands nur bis zu einem gewissen Grenzwert von V_{GS} erreicht, wogegen sich oberhalb dieses Grenzwerts der lineare Bereich ohne ein Erreichen von $V_{DS,sat}$ für alle V_{DS} fortsetzt.

8.3 Vergleich der untersuchten FETs

Die Unterschwelkenennlinien von FET 1 und FET R zeigen qualitativ einen sehr ähnlichen Verlauf. Zu beachten ist hierbei aber, dass FET R bei kleineren *Gate*-Spannungen gemessen wurde und dass dieser Transistor ein dünneres Siliziumdioxid aufweist. Für einen quantitativen Vergleich muss daher die Unterschwelkenennlinie von FET 1 zunächst im kleineren Bereich der *Gate*-Spannung V_{GS} von -5 bis $+5$ V ausgewertet werden.

Für einen vollständigen Vergleich muss dann auf die Kapazität des *Gate*-Kondensators normiert werden. Hierfür ist der relative Unterschied der Schichtdicke des Siliziumdioxids und der effektiven Kanalfäche zwischen FET 1 und FET R nötig. Für den Wafer, auf dem FET R gemessen wurde, konnte in [6] ermittelt werden, dass der Kanal in Transistoren von diesem Typ von nur einer CNT gebildet wird. Für FET 1 ist die CNT-Anzahl wegen der weiter unten diskutierten Schwierigkeiten hinsichtlich der effektiven Kanallänge und -weite nicht verfügbar. Um dennoch eine grundsätzliche Vergleichbarkeit zu ermöglichen, wird hier lediglich auf die Schichtdicke des Siliziumdioxids normiert. Der Unterschied der Schichtdicken ist bekannt und führt dann zu einer 2,5fach höheren Kapazität des *Gate*-Kondensators von FET R.

Wie bereits in Abb. 59a notiert, weist FET 1 bei einer *Gate*-Spannung von -5 V einen *On*-Strom von 800 nA auf. Normiert auf die dünnere Oxidschicht entspricht dies einem *On*-Strom von $1,975\text{ }\mu\text{A}$, was um rund einen Faktor 3 höher ist als der *On*-Strom von FET R. Die Frage, ob dieser höhere Strom von FET 1 durch drei CNTs verursacht wird, oder ob es eine CNT ist, die bessere Eigenschaften als die CNT in FET R aufweist, kann hier nicht abschließend geklärt werden.

Effektive Kanallänge und -weite

Der hier genutzte Prozess zur Herstellung von CNTFETs definiert nur einen Kanalbereich, in dem die CNTs gewachsen werden. Es findet dabei nur eine Justage der Kontakte auf die Kanalbereiche, aber nicht auf die CNTs selbst, statt. Die gewachsenen CNTs sind somit weder senkrecht zwischen den *Source/Drain*-Kontakten ausgerichtet noch weisen sie eine gemeinsame Orientierung auf. Es kommt also zu einem Unterschied zwischen den nominalen, durch das Testchip-Layout definierten, und den effektiven Abmessungen des Kanals. Für die hier betrachteten Fingertransistoren ist der Zusammenhang zwischen der nominalen und effektiven Kanallänge bzw. -weite über Gl. 45 bis Gl. 47 gegeben und in Abb. 60 skizziert.

$$W_{\text{Kanal|eff}} = \sum_i d_{\text{SWNT},i} \quad (45)$$

$$L_{\text{Kanal|eff,min}} = L_{\text{Kanal|nom}} \quad (46)$$

$$L_{\text{Kanal|eff,max}} = \sqrt{L_{\text{Kanal|nom}}^2 + 180\text{ }\mu\text{m}^2} \quad (47)$$

Die effektive Kanalweite $W_{\text{Kanal|eff}}$ wird dabei über den Durchmesser d_{SWNT} der CNT bestimmt, die ihn bildet. Sollte es mehr als eine *Source/Drain*-Verbindung geben, entspricht die effektive Kanalweite der Summe der jeweiligen CNT-Durchmesser. Die effektive Kanallänge $L_{\text{Kanal|eff}}$ ist von der konkreten Position der CNT zwischen *Source* und *Drain* abhängig. Der minimale Wert $L_{\text{Kanal|eff,min}}$, der erreicht wird, wenn die CNT exakt senkrecht zwischen *Source* und *Drain* liegt, entspricht dabei der nominalen Länge $L_{\text{Kanal|nom}}$ des Kanals. Der maximale Wert $L_{\text{Kanal|eff,max}}$ wird erreicht, wenn die CNT diagonal im Kanalbereich liegt. Hierbei ist zu beachten, dass durch die Fingerstruktur der Transistoren mit n Fingerpaaren die Kanalweite $W_{\text{Kanal|nom}}$ in einzelne Segmente von $180\text{ }\mu\text{m}$ Weite aufgeteilt wird; vgl. auch Abb. 37 (Seite 59).

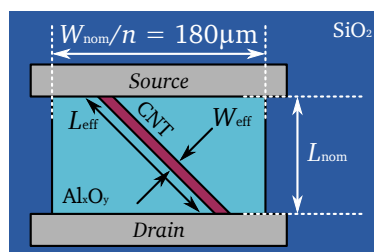


Abb. 60: Nominale und effektive Abmessungen des Kanals in einem CNTFET.

Tab. 6: Vergleich von nominaler und effektiver Kanallänge bzw. -weite und nominale Fläche des Kanalbereichs.

	$L_{\text{Kanal nom}}$	$W_{\text{Kanal nom}}$	Kanalgeometrie	$A_{\text{Kanal nom}}$	$L_{\text{Kanal eff,min}}$	$L_{\text{Kanal eff,max}}$
FET 1	5 μm	5760 μm	5 $\mu\text{m} \times (32 \cdot 180 \mu\text{m})$	28 800 μm^2	5 μm	900 μm
FET 2	2 μm	5760 μm	2 $\mu\text{m} \times (32 \cdot 180 \mu\text{m})$	11 520 μm^2	2 μm	360 μm
FET R	5 μm	5 μm	5 $\mu\text{m} \times 5 \mu\text{m}$	25 μm^2	5 μm	7 μm

Für die drei in diesem Kapitel betrachteten Transistoren sind die entsprechenden Werte samt ihrer Kanalgeometrie und der nominalen Fläche $A_{\text{Kanal|nom}}$ des Kanalbereichs in Tab. 6 zusammengestellt. Wie Tab. 6 zeigt, kann die effektive Kanallänge für FET 1 und FET 2 aufgrund der großen nominalen Kanalweiten in einem weiten Bereich schwanken. Dass der Wert von $L_{\text{Kanal|eff,max}}$ jemals erreicht wird, ist unwahrscheinlich, da die hier bisher beobachteten CNTs eine Länge von einigen Mikrometern nicht überschreiten. Da aber gleichzeitig die nominale Fläche des Kanalbereichs größer ist, als dass sie mittels AFM mit vertretbarem Aufwand vermessen werden könnte, besteht hier hinsichtlich der effektiven Abmessungen des Kanals eine deutliche Unsicherheit.

Zusammenhang zwischen Schaltverhalten und nominaler Kanallänge

Die im letzten Unterkapitel vorgestellten repräsentativen Messungen zeigen, dass FET 1 ein gutes Schaltverhalten und FET 2 ein schlechtes Schaltverhalten aufweist. FET 1 hat dabei eine nominale Kanallänge von 5 μm und FET 2 von 2 μm . Auf Basis der nominalen Kanallängen kann ein empirischer Zusammenhang zwischen dem Schaltverhalten und der Kanallänge abgeleitet werden. Dieser Zusammenhang kann wegen des großen Messaufwands nicht analytisch verifiziert werden; obwohl er also hypothetisch bleibt, beschreibt er den Einfluss der nominalen Kanallänge auf das Schaltverhalten in plausibler Weise.

Unter Beachtung der in Kapitel 7.7 (Seite 83) präsentierten Ergebnisse, dass die durchschnittliche Länge einer CNT im Bereich von 2 bis 3 μm liegt, kann das schlechte Schaltverhalten von FET 2 zunächst wie folgt erklärt werden: Da die nominale Kanallänge im Bereich der durchschnittlichen CNT-Länge liegt und aufgrund der enormen nominalen Kanalweite von 5760 μm , kann davon ausgegangen werden, dass nicht nur eine, sondern eine Vielzahl an *Source/Drain*-Verbindungen durch CNTs zustande kommt. Der Kanal von FET 2 wird also durch viele verschiedene CNTs gebildet, die sowohl metallisch als auch halbleitend sein können. In einer solchen Parallelschaltung wird das Verhalten von denjenigen Verbindungen bestimmt, die den geringsten elektrischen Widerstand aufweisen; d. h. von metallischen CNTs und halbleitenden CNTs mit nur kleiner Bandlücke, woraus für das gesamte Bauelement ein schlechtes Schaltverhalten resultiert.

In analoger Weise kann das gute Schaltverhalten von FET 1 erklärt werden: Dadurch dass die nominale Kanallänge größer ist als die typische CNT-Länge, ist es trotz der großen nomina-

len Kanalweite von ebenfalls $5760\text{ }\mu\text{m}$ unwahrscheinlich, dass viele *Source/Drain*-Verbindungen zustande kommen. Damit steigt für eventuell vorhandene halbleitende Verbindungen die Wahrscheinlichkeit, nicht durch metallische Verbindungen kurzgeschlossen zu werden. Zusätzlich kann hier eine Kettenbildung von CNTs zum Tragen kommen [86]. Eine leitfähige Verbindung zwischen *Source* und *Drain* wird dabei nicht durch eine einzelne, sondern durch mehrere sich berührende CNTs gebildet. So entsteht eine *Source/Drain*-Verbindung, die aus einer Mischung von CNTs zusammengesetzt ist, d. h. sie besteht aus einer Serienschaltung von sowohl halbleitenden als auch metallischen CNTs. Eine solche CNT-Kette verhält sich halbleitend, sobald eines ihrer Serienelemente eine halbleitende CNT ist, da sich der Gesamtwiderstand aus der Summe der Einzelwiderstände zusammensetzt und damit über den Feldeffekt steuerbar wird. Da hierfür nur ein einziges Element halbleitend sein muss, nimmt die Wahrscheinlichkeit für ein halbleitendes Verhalten der gesamten Kette mit der Anzahl der Elemente bzw. der Länge der Kette zu.

9 Untersuchung des dielektrischen Versagens der CNTFETs mit Al/Ni-Katalysator

Dieses Kapitel ist der Suche nach der Ursache für die im vorherigen Kapitel festgestellte äußerst geringe Ausbeute an funktionalen FETs aufgrund dielektrischen Versagens des *Gate*-Oxids gewidmet. Ziel ist es, den Ort der Durchbrüche zu lokalisieren und den Mechanismus der Schädigung zu ermitteln, um auf dieser Basis den Herstellungsprozess für CNTFETs so zu modifizieren, dass elektrisch stabile *Gate*-Dielektrika hergestellt werden.

9.1 Lokalisierung der dielektrischen Durchbrüche

Um die dielektrischen Durchbrüche zu lokalisieren, wurde an einem Bauelement, das dielektrisches Versagen gezeigt hat, eine Messung mittels Strom-Spannungs-Mikroskopie (CS-AFM) durchgeführt. Diese Methode ist geeignet, orts aufgelöst Schäden in Dielektrika zu erfassen [87, 88]. Da die *Source/Drain*-Kontakte des FETs aber zunächst eine orts aufgelöste Untersuchung der Oxidschäden verhindern, indem sie durch ihre gute elektrische Leitfähigkeit sämtliche möglicherweise von einander isolierten Schäden kurzschließen, wurden sie nasschemisch mit einer modifizierten Königswasserlösung entfernt. Die verwendete Lösung entfernt nicht nur Palladium, sondern auch alle anderen hier genutzten Metalle, sodass auch eventuelle metallische Reste des Katalysators in Form von Nickel und Aluminium entfernt wurden. Siliziumdioxid und Aluminiumoxid werden dagegen ebenso wie Kohlenstoff nicht angegriffen.

Für die CS-AFM-Messung wurde eine Spannung von 5 V zwischen Probe und Messspitze angelegt und als Messstelle eine Ecke einer Katalysator-Insel gewählt. Das Strom- und das Höhen-Signal der Messung sind in Abb. 61a und Abb. 61b dargestellt, während Abb. 61c ein Profil entlang der gestrichelten Linie beider Signale zeigt; die violette Markierung zeigt korrespondierende Bereiche. Das Strom-Signal der Messung zeigt für große Bereich der vermessenen Fläche den gleichen Strom von 0,35 pA. Insbesondere gilt dies sowohl für die Fläche des Oxids, auf der sich zu keiner Zeit des Prozesses Metall befunden hat, als auch für den Bereich der Katalysator-Inseln, in dem das Metall nasschemisch entfernt wurde. Allein die scharf umrissene Kante der Katalysator-Insel zeigt einen erhöhten Strom von 0,45 pA, womit die Bruchstellen des Oxids eindeutig auf diese Randbereiche lokalisiert werden können. Dass das Strom-Signal die Kante

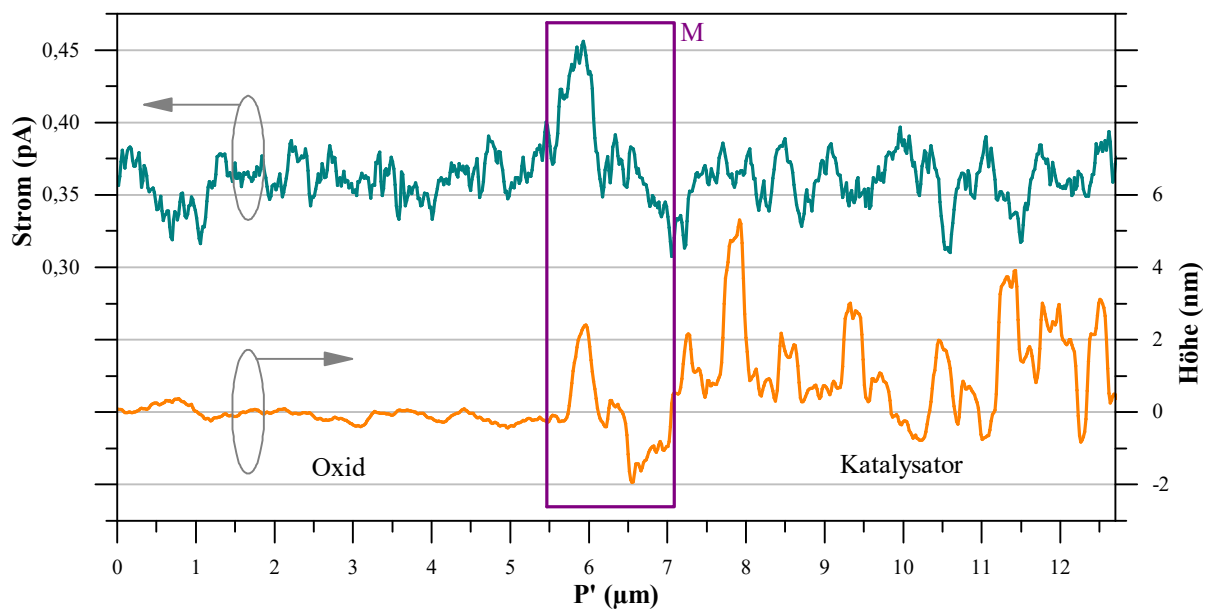
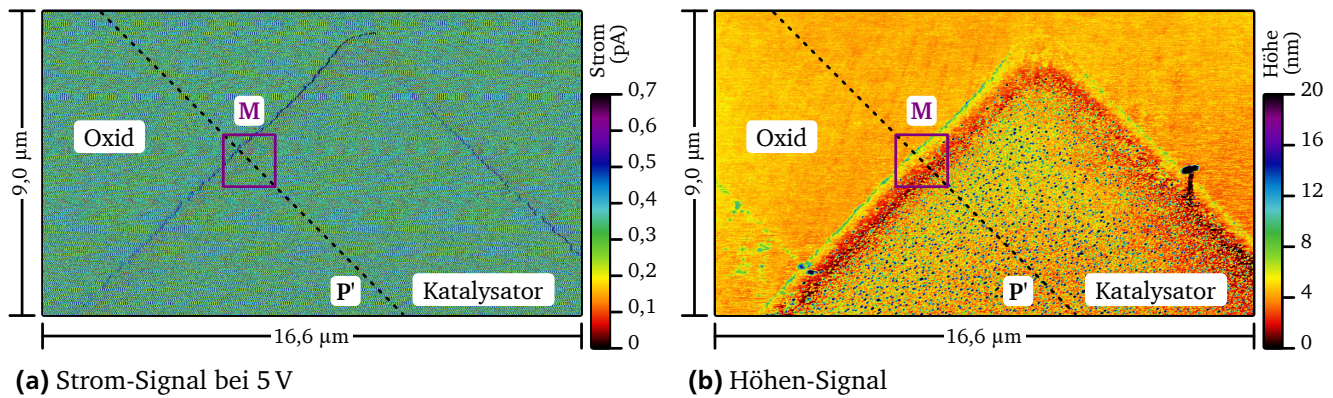


Abb. 61: CS-AFM-Messung zur Lokalisierung der dielektrischen Durchbrüche auf einer zuvor nasschemisch von allen Metallen befreiten Probe.

der Katalysator-Insel auf der rechten Seite leicht nach links versetzt abbildet, wird durch eine einseitige Abnutzung der leitfähigen Beschichtung der Messspitze verursacht und ist damit ein Messartefakt.

Das Höhen-Signal der Messung in Abb. 61b zeigt für den Bereich des Oxids eine glatte und für den Bereich des Katalysators eine mit Clustern bedeckte Oberfläche. Dass die aus Nickel bestehenden Cluster nicht nasschemisch entfernt wurden, kann mit einem Belag aus Kohlenstoff erklärt werden, der das Metall vor dem nasschemischen Angriff schützt: Da es sich um einen vollständig prozessierten Wafer handelt, hat dieser auch den CCVD-Prozess erfahren, weswegen die Cluster dem in Kapitel 7.1 (Seite 67) dargestellten CNT-Wachstumsmechanismus entsprechend mit Kohlenstoff umhüllt sind. Dass eine Einhüllung der Cluster mit Kohlenstoff tatsächlich vorliegt, zeigt auch die TEM-Messung in Abb. 62.

9.2 Mechanismus der Durchbrüche

Die in Abb. 62 gezeigte TEM-Messung wurde im Rahmen einer anderen Arbeit am Institut als externe Auftragsarbeit angefertigt [7, 89]. Die im Querschnitt vermessene Probe wurde dabei mit leicht anderen Parametern hergestellt: Der Katalysator besteht aus dickeren Schichten, nämlich 10 nm Aluminium und 10 nm Nickel, während das thermisch gewachsene Siliziumdioxid dagegen mit $d_{\text{Oxid|initial}} = 50$ nm dünner ist. Auch wenn quantitative Rückschlüsse wegen der anderen Schichtdicken nicht möglich sind, kann diese Messung Aufschluss über den Mechanismus der Schädigung geben.

Die eigentliche Messung in Abb. 62a ist in Abb. 62b schematisch nachgezeichnet. Auf der Messung sind die folgenden Strukturen zu erkennen: eine CNT im Querschnitt; ein Cluster, der von einem dünnen Film bedeckt ist; eine Schicht mit rauer Oberfläche, die von dem gleichen Film überzogen ist wie der Cluster; und schließlich drei weitere Cluster in dieser Schicht, die zum Teil in das darunter liegende Siliziumdioxid „einsinken“. Die entscheidende Erkenntnis ist, dass es offensichtlich zwei Arten von Clustern gibt: solche, die auf der Oberfläche liegen und solche, die in das Oxid eingebettet sind.

Dass neben den Nickel-Clustern eine weitere Art von Clustern gebildet wird, kann wie folgt erklärt werden: Während der Temperung bei 920 °C wird das Aluminium, das einen Schmelzpunkt von 660 °C hat, flüssig. Oberhalb von 450 °C kommt es zwischen Aluminium und Siliziumdioxid zu einer Redoxreaktion nach Gl. 48 [90].

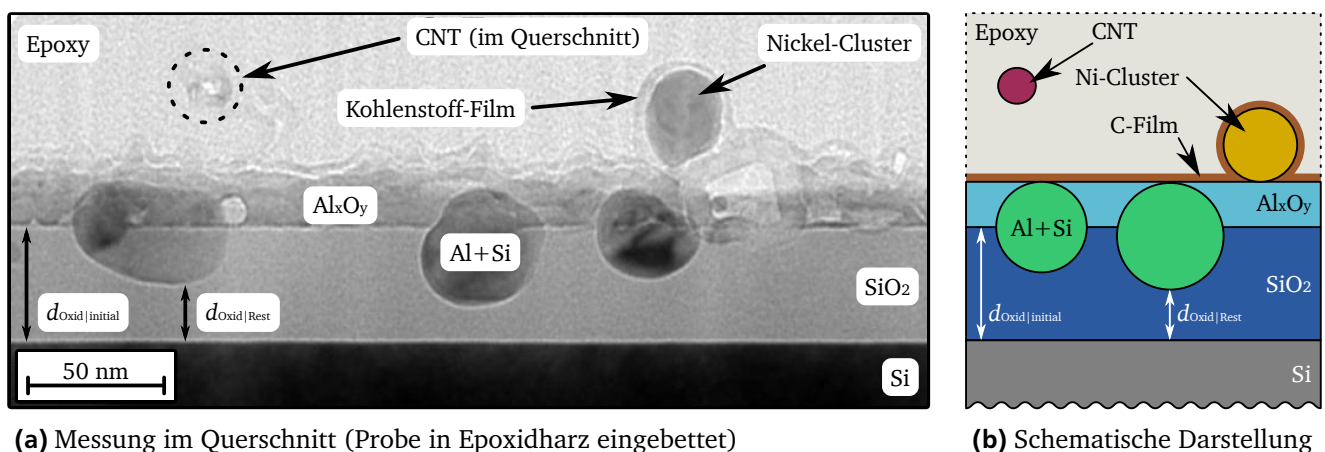


Abb. 62: TEM-Messung, die im Rahmen einer anderen Arbeit am IHTN als Auftragsarbeit angefertigt wurde [7, 89]: Die gemessene Probe wurde mit leicht anderen Parametern hergestellt: 10 nm Aluminium und 10 nm Nickel als Katalysator auf 50 nm thermisch gewachsenem Siliziumdioxid. Trotz der anderen Parameter steht diese Messung qualitativ auch für den in dieser Arbeit genutzten Katalysator.

Aufgrund der hohen Beweglichkeit und Löslichkeit von Aluminium in Siliziumdioxid kommt es allerdings auch zum Transport von Aluminium in tiefere Bereiche des Oxids und zu einer Legierungsbildung [91, 92]. Die Legierungsbildung geschieht in Form der tieferliegenden Al + Si-Cluster, die in das Oxid einsinken. Während an der Aluminium–Siliziumdioxid-Grenzfläche diese Legierungsbildung stattfindet, wird das Nickel auf der Oberfläche ebenfalls zu Clustern „zusammengeschoben“. Nickel hat als Volumenmaterial einen Schmelzpunkt von 1455 °C und sollte damit während des gesamten Prozesses in fester Form vorliegen. Da Schmelzpunkte für feine Partikel und dünne Schichten aber niedriger sind als die des Volumenmaterials, kann hier davon ausgegangen werden, dass das Nickel leicht duktil wird und somit eine „Umformung“ ermöglicht. Durch den CCVD-Prozess kommt es nach der Temperung zur Aufnahme von Kohlenstoff an den Nickel-Clustern. Dass die gesamte Oberfläche der hier gezeigten Probe mit Kohlenstoff bedeckt ist, liegt in den auch während des CCVD-Prozesses abweichenden Parametern der Herstellung begründet: In [7] wurde mit der vierfachen Menge Methan gearbeitet und einer dickeren Schicht Nickel gearbeitet, da das Wachstum von Graphen und nicht das von CNTs angestrebt wurde.

Die Bildung der Al + Si-Cluster und deren Einsinken in das Siliziumdioxid führt zu einer lokalen Reduzierung der Oxiddicke, einer Art „Oxid-Verbrauch“. Im Bereich der TEM-Messung in Abb. 62 findet eine Reduzierung vom ursprünglichen Wert der Dicke von $d_{\text{Oxid|initial}} = 50 \text{ nm}$ auf eine Restdicke $d_{\text{Oxid|Rest}}$ von nur noch rund 25 nm statt. Dieser Oxid-Verbrauch ist hier damit für das dielektrische Versagen der CNTFETs mit Aluminium/Nickel-Katalysator verantwortlich zu machen: Durch die *Source/Drain*-Kontakte werden im Bereich eines FETs alle lokalen Schwächungen des Oxids elektrisch kurzgeschlossen, sodass für das ganze Bauelement nur eine möglicherweise signifikant reduzierte Oxiddicke wirksam ist. Erschwerend kommt hinzu, dass hierbei der im Bereich des FETs größte Al + Si-Cluster ausschlaggebend ist, also derjenige Cluster, der zur insgesamt kleinsten Restdicke $d_{\text{Oxid|Rest}}$ des Oxids führt.

9.3 Ursache des Versagens

Mit Hilfe der CS-AFM- und der TEM-Messungen konnten in den letzten beiden Unterkapiteln zwei Dinge festgestellt werden: Die dielektrischen Durchbrüche sind in der Hauptsache auf die Kanten der Katalysator-Inseln beschränkt und der Mechanismus der Durchbrüche basiert auf einer lokalen Reduzierung der Oxiddicke durch die Bildung von Al + Si-Clustern („Oxid-Verbrauch“). Zunächst noch unklar ist das generelle Auftreten der Durchbrüche: Die in Abb. 62 gezeigten Al + Si-Cluster führen zwar zu einer Reduzierung der Oxiddicke um rund 25 nm, da die in dieser Arbeit verwendete Schichtdicke für Aluminium mit 5 nm aber niedriger ist als die der gezeigten Probe, sollten wegen des geringeren Angebots an Aluminium auch die

Abmessungen der Al + Si-Cluster geringer sein. Und selbst wenn es zu einer Reduzierung der hier initial vorhandenen 100 nm Oxid um 25 nm auf 75 nm Restdicke kommen sollte, sollte diese Oxiddicke für die in Kapitel 8.2 gewählten *Gate-Spannung* V_{GS} von ± 10 V ausreichend sein; in der Vorgängerarbeit konnten immerhin FETs bei *Gate-Spannungen* V_{GS} von ± 5 V auf bis zu 30 nm dünnen Oxiden problemlos vermessen werden.

Der wesentliche technologische Unterschied zwischen dieser und der Vorgängerarbeit besteht in der Einführung des strukturierten Katalysators. Deswegen und weil die Durchbrüche auf die Kanten der Katalysator-Inseln beschränkt sind, liegt der Verdacht nahe, dass wegen der Strukturierung des Katalysators an den Kanten der Inseln mehr Aluminium für die Bildung der schädigenden Al + Si-Cluster zur Verfügung steht. Die Ursache für dort zusätzlich vorhandenes Aluminium ist ein Artefakt der *Lift off*-Technik: das sogenannte *Fencing*, dessen Problematik in Abb. 63 skizziert ist. *Fencing* entsteht, wenn nach dem Entwickeln während der Schichtabscheidung nicht nur die Oberfläche des Substrats wie in Abb. 63a gezeigt, sondern auch die unteren Teile des Lacks beschichtet werden. Diese seitliche Beschichtung übersteht, je nach Schichtdicke und Kontakt zur eigentlichen Schicht, auch den *Lift off*-Prozess und erzeugt um die Strukturen herum einen Überstand an der Kante, der wie in Abb. 63b dargestellt, wie ein „Zaun“ aussieht. Die Höhe und Breite des *Fencings* kann dabei ein Vielfaches der eigentlichen Schichtdicke, die strukturiert werden soll, übersteigen. Abb. 63c zeigt eine AFM-Messung an der Kante einer Katalysator-Insel und wie das Profil in Abb. 63d deutlich macht, ist das *Fencing* mit 30 nm rund sechsmal so hoch wie die eigentliche Höhe des Katalysators von 5 nm.

Erschwerend zum generellen Auftreten von *Fencing* kommt hier hinzu, dass hinsichtlich der Prozessierung vermeintlich geringfügige Änderungen vorgenommen wurden: Zunächst wurde bei der Elektronenstrahlverdampfung, die zur Abscheidung des Katalysators verwendet wird, der Probenhalter von einer stationären Halterung für zwei Wafer auf ein Planetengetriebe umgestellt, das bis zu 30 Wafer fasst. Dies wurde mit dem Ziel gemacht, die Schwankungen zwischen den Prozessen zu umgehen, die bedingt durch die Konstruktion der Anlage nicht

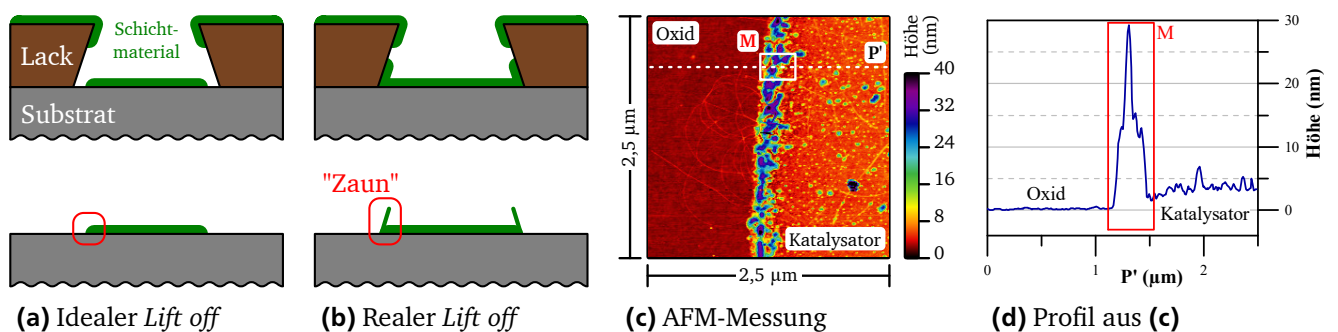


Abb. 63: Nicht idealer *Lift off* mit *Fencing*-Problem: Durch eine teilweise Beschichtung der Lackflanken entsteht um die Strukturen herum *Fencing*, das in Form eines Überstands an der Kante in AFM-Messungen sichtbar wird und ein Vielfaches der eigentlichen Schichtdicke erreichen kann.

gänzlich vermieden werden können. Durch die Verwendung des Planetengetriebes ist die Abscheidung allerdings konformer und weniger gerichtet, was bedeutet, dass auch die Flanken des *Lift off*-Lacks teilweise beschichtet werden. Außerdem wurde im Rahmen der Lithographie das Lacksystem gewechselt. Zunächst wurde für den *Lift off* mit einem Doppellagen-System gearbeitet, das sich aber als extrem anfällig gegen Schwankungen in der Raumtemperatur und Luftfeuchte gezeigt hat. Aufgrund dieser Unzuverlässigkeit mussten mehrfach Wafer im letzten Arbeitsschritt verworfen werden, da der *Lift off* gescheitert und die Proben damit nicht ordnungsgemäß metallisiert waren und nicht elektrisch charakterisiert werden konnten. Der Unterschied zwischen dem alten Lacksystem und dem neuen System mit Umkehrlack besteht in der Form der Flanken: der Umkehrlack erzeugt, wie auch in Abb. 63a nochmals skizziert, V-förmige Flanken, während das Doppellagensystem zu T-förmigen Flanken führt. Dadurch dass eine T-förmige Flanke im unteren Bereich steiler ist und vom oberen Bereich des Lacks abgeschattet wird, zeigt ein Doppellagensystem eine geringere Anfälligkeit für *Fencing*.

Das *Fencing*, das hier als Artefakt der Strukturierung des Katalysators entsteht, muss also für das Auftreten der dielektrischen Durchbrüche an den Kanten der Katalysator-Inseln verantwortlich gemacht werden. Die deutlich größeren Mengen an Aluminium führen während der folgenden Temperung zu großen Al + Si-Clustern, die die Oxiddicke lokal stark reduzieren. Diese lokal verringerte Oxiddicke führt dann schließlich für das gesamte Bauelement zu einer deutlich verringerten Durchbruchspannung.

10 Evaluierung anderer Katalysatoren

In Kapitel 9 konnte ermittelt werden, dass für das dielektrische Versagen der CNTFETs mit Aluminium/Nickel-Katalysator das sogenannte *Fencing* verantwortlich ist, das ein Artefakt der Strukturierung des Katalysators mittels *Lift off* ist. In Verbindung mit dem inhärenten Oxid-Verbrauch des Aluminium/Nickel-Katalysators ist *Fencing* der Grund für die äußerst geringe Ausbeute an funktionalen FETs. Der Oxid-Verbrauch ist dabei darauf zurückzuführen, dass ein Teil des Aluminiums während der Temperung nicht vollständig zu Aluminiumoxid transformiert wird, sondern eine leitfähige Legierung mit dem Silizium bildet. Diese Al + Si-Cluster „durchdringen“ wie nochmals in Abb. 64a skizziert das Dielektrikum und verursachen auf diese Weise eine lokale Dünnung des Dielektrikums und die in Kapitel 8 aufgetretenen *Gate*-Durchbrüche. Die Bildung der Al + Si-Cluster kann nie völlig ausgeschlossen werden, da die Redoxreaktion nach Gl. 48 (Seite 95), die die Transformation des metallischen Aluminiums zum dielektrischen Aluminiumoxid beschreibt, immer auch Silizium erzeugt — das dann seinerseits als Legierungspartner für weiteres metallisches Aluminium zur Verfügung steht. Dass die Durchbrüche besonders in den Randbereichen der Katalysator-Inseln verstärkt zu finden sind, liegt darin begründet, dass dort durch *Fencing* zusätzliches Aluminium vorhanden ist. Um eine Schädigung des Oxids durch die Bildung von Al + Si-Clustern zu vermeiden, sind folgende Ansätze denkbar, die hier kurz diskutiert werden sollen:

- Der Ansatz des „sacraficial catalyst“, der von der Vorgängerarbeit verfolgt wird [6, 93], zeigt nur einen geringen Oxid-Verbrauch. Da der genutzte Aluminium/Nickel-Katalysator nicht strukturiert wird, werden keine lokalen Ansammlungen von metallischem Aluminium erzeugt, sodass nur moderat Al + Si-Cluster auftreten.
- Die Verwendung eines direkten Lithographie-Prozesses zur Strukturierung des Katalysators anstelle des *Lift off*-Prozesses würde *Fencing*, und damit auch überschüssiges Aluminium, prinzipiell verhindern. Die Umsetzung dieser Möglichkeit gestaltet sich allerdings schwierig, da die nasschemische Strukturierung von Aluminium an sich und besonders bei gleichzeitiger Gegenwart von Nickel aufwendig und in den Ergebnissen nicht zufriedenstellend ist; daher wird dieser Ansatz hier nicht weiter verfolgt.
- Ein anderer in Abb. 64b skizzierter Ansatz ist die Einführung einer dielektrischen „Barriere“ unterhalb des Siliziumdioxids. Die Barriere hat den Zweck, das Einsinken der sich bildenden Al + Si-Cluster zu stoppen bzw. so stark zu reduzieren, dass keine *Gate*-Durchbrüche auftreten. Das Siliziumdioxid selbst kann dann mit einer geringeren Schichtdicke ausge-

führt werden, da es nur noch die Aufgabe hat, auf der Probenoberfläche den für die Cluster-Bildung bis hierhin gewohnten Untergrund zur Verfügung zu stellen. Dieser Ansatz geht damit davon aus, dass der Mechanismus der Cluster-Bildung konkret auf die zwischen Aluminium und Siliziumdioxid ablaufende Reaktion angewiesen ist.

- Ein weiterer Ansatz ist in Abb. 64c skizziert. Dieser Ansatz geht davon aus, dass die Cluster-Bildung nicht direkt auf die Transformation von Aluminium zu Aluminiumoxid selbst angewiesen ist, sondern nur auf einen dielektrischen Untergrund. Da in diesem Fall die Redoxreaktion von Aluminium mit Siliziumdioxid ausbliebe, wäre auch die Bildung der Al + Si-Cluster unterbunden. Der Prozess würde dann dementsprechend modifiziert, dass für den Katalysator nicht ein metallischer Doppelstapel mit Aluminium und Nickel abgeschieden würde, sondern eine Nickel-Schicht auf einer dielektrischen Schicht.

10.1 Schichtstapel der Katalysatoren

Um zu klären, ob mit den beiden neuen in Abb. 64 skizzierten Ansätzen überhaupt CNTs gewachsen werden, werden vier verschiedene Schichtstapel untersucht. Stapel 1 verfolgt dabei den ersten Ansatz mit einer „Barriere“-Schicht, während Stapel 2 bis Stapel 4 den zweiten Ansatz mit alternativen Dielektrika nutzen. Die jeweilige Zusammensetzung der Schichtstapel ist in Tab. 7 zusammengefasst. Als Substrate werden Wafer mit 100 nm Trockenoxid verwendet.

Für Stapel 1 werden auf die vorbereiteten Substrate mittels CVD und des Silan/Ammoniak-Prozesses zunächst 20 nm Siliziumnitrid aufgewachsen. Durch eine Feuchtoxidation wird im

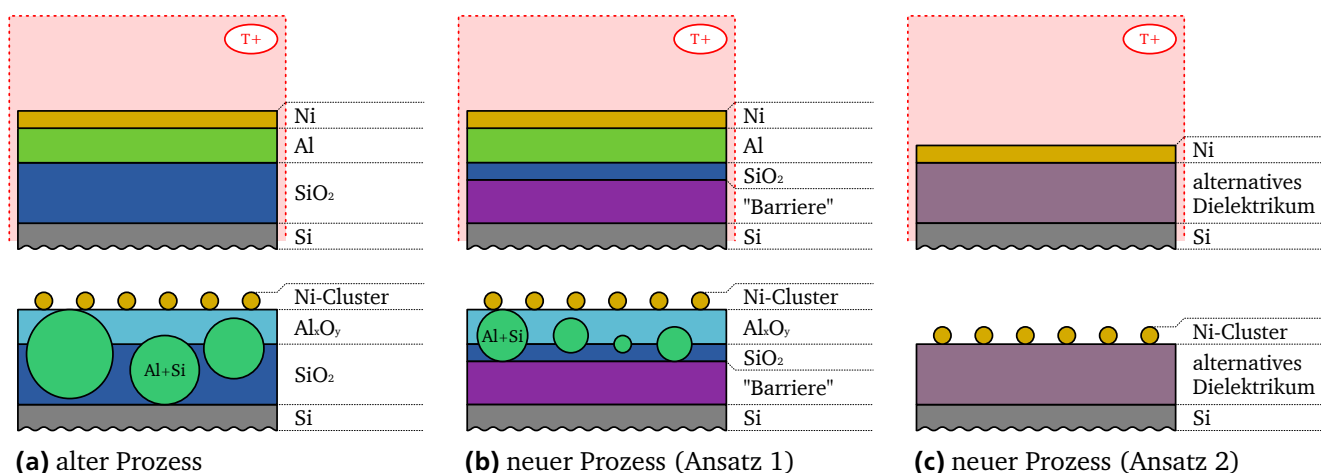


Abb. 64: Alternative Prozess-Konzepte zur Vermeidung der *Gate*-Durchbrüche: Durch die Einführung einer „Barriere“ **(b)** bzw. alternativer Dielektrika **(c)** sollen *Gate*-Durchbrüche verhindert werden, indem die Bildung der Al + Si-Cluster eingeschränkt bzw. von vornherein vermieden wird. Für **(a)** vgl. Abb. 62.

Tab. 7: Alternativen Schichtstapel samt Abscheidemethoden des Dielektrikums, wobei Stapel 1 den Ansatz mit „Barriere“-Schicht (vgl. Abb. 64b) verfolgt, während Stapel 2 bis Stapel 4 den Ansatz der alternativen Dielektrika (vgl. Abb. 64c) nutzen. Als Substrate werden Wafer mit 100 nm Trockenoxid verwendet.

	Katalysator	Dielektrikum	Abscheidemethode
Stapel 1 (Ansatz 1)	Aluminium & Nickel (5 nm & 1 nm)	Siliziumnitrid (10 nm, Si_3N_4)	chemische Gasphasenabscheidung vgl. Kapitel 5.7 (Seite 44)
		Siliziumdioxid (10 nm, SiO_2)	Nitrid-Umwandlung durch Feuchtoxidation vgl. Kapitel 5.2 (Seite 35)
Stapel 2 (Ansatz 2)	Nickel (1 nm)	Aluminiumoxid (5 nm, Al_2O_3)	reaktives <i>Sputtern</i> (Prozess nach [94]) vgl. Kapitel 5.9 (Seite 50)
Stapel 3 (Ansatz 2)	Nickel (1 nm)	Aluminiumnitrid (5 nm, AlN)	reaktives <i>Sputtern</i> (Prozess nach [95]) vgl. Kapitel 5.9 (Seite 50)
Stapel 4 (Ansatz 2)	Nickel (1 nm)	Aluminiumoxid (5 nm, Al_2O_3)	Atomlagenabscheidung (TMA/Wasser-Prozess) vgl. Kapitel 5.8 (Seite 46)

Anschluss die obere Hälfte des Nitrids oxidiert, was zu einer 10 nm dicken Oxidschicht auf dem Nitrid führt. Auf diesen Nitrid–Oxid-Stapel werden schließlich die bisher gewohnten 5 nm Aluminium und 1 nm Nickel mittels Elektronenstrahlverdampfung aufgebracht. Auf Stapel 2 werden 5 nm Aluminiumoxid mit Hilfe eines reaktiven *Sputter*-Prozesses nach [94] abgeschieden. Gesputtert wird mit Argon aus einem reinen Aluminium-*Target*, wobei dem Argon im Verhältnis 1 : 150 Sauerstoff beigemischt wird. Zusätzlich wird hier eine „Nachoxidation“ in Ozon für 30 min mit UV-Unterstützung durchgeführt, die metallisches Aluminium in der Schicht vermeiden soll [96]. Diese Nachoxidation geschieht mittels des in Kapitel 5.12 (Seite 53) beschriebenen Eigenbaus.

Auf Stapel 3 werden in analoger Weise 5 nm Aluminiumnitrid nach [95] abgeschieden, wobei Stickstoff anstelle von Sauerstoff als Reaktivgas verwendet wird. Für Stapel 4 wird mittels ALD und des TMA/Wasser-Prozesses Aluminiumoxid auf die Substrate abgeschieden. Um die angestrebte Schichtdicke von 5 nm zu erreichen, werden dabei 60 Zyklen benötigt. Abschließend werden auf diese drei Schichtstapel jeweils 1 nm Nickel mittels Elektronenstrahlverdampfung aufgebracht.

10.2 CNT-Wachstum auf den untersuchten Katalysatoren

Da über die Details der Cluster-Bildung und des CNT-Wachstums mit den alternativen Schichtstapeln wenig bekannt ist, soll zunächst mit einer möglichst einfachen Prozessführung gearbeitet werden, die deswegen einige der Erkenntnisse aus Kapitel 7 (Seite 67) ignoriert. Es wird hier daher der Basisprozess gemäß Abb. 47 (Seite 70) mit einer Cluster-Bildungs-Temperatur T_{Temp} und einer CNT-Wachstums-Temperatur T_{CCVD} von jeweils 870 °C mit Argon als Inertgas verwendet.

Die Ergebnisse des CNT-Wachstums auf den vier alternativen Schichtstapeln zeigt Abb. 65. Die AFM-Messungen lassen erkennen, dass auf allen vier Stapeln Cluster gebildet und CNTs gewachsen werden. Für den Mechanismus der Cluster-Bildung bedeutet dies, dass ein Ablaufen der Redoxreaktion von Siliziumdioxid und Aluminium direkt unter dem Nickel nicht zwingend erforderlich ist, sondern ein dielektrischer Untergrund für das Nickel ausreichend ist. Was das Wachstum der CNTs angeht, sind die hier gewählten Parameter der gemeinsamen und vereinfachten Prozessführung für alle Schichtstapel ausreichend passend; es dürfte aber für jeden Schichtstapel noch Potential für Optimierungen geben, sodass eine eingehendere quantitative Bewertung dieser Ergebnisse zu diesem Zeitpunkt wenig Sinn ergibt. Hier kann nun nur festgehalten werden, dass für den Basisprozess das beste Ergebnis hinsichtlich Länge, Anzahl und Durchmesser der CNTs für Stapel 4 erzielt wird.

Obwohl auf allen alternativen Schichtstapeln CNTs gewachsen werden, kann hier nur Stapel 4 als mögliche Alternative für den bisher genutzten Aluminium/Nickel-Katalysator genannt werden. Sowohl der auf einem CVD-Prozess basierende Stapel 1 als auch die beiden auf einem reaktiven *Sputtern* basierenden Stapel 2 und Stapel 3 zeigen hinsichtlich ihrer Prozessierbarkeit Unzulänglichkeiten. Auf Stapel 1 treten nach der Abscheidung des Siliziumnitrids in einigen Fällen Schlammflecken-artige Strukturen auf, die das CNT-Wachstum dann verhindern. Für die beiden mittels reaktivem *Sputtern* abgeschiedenen Schichten war es entgegen aller Erwartungen nicht möglich, die Schichten eindeutig als Aluminiumoxid bzw. -nitrid zu identifizieren. Hierfür wurde jeweils die Schichtdicke mittels AFM bestimmt und dann unter Annahme dieser Schichtdicke mittels Ellipsometrie der Brechungsindex bestimmt — wobei keine Übereinstimmung mit tabellierten Werten gefunden werden konnte. Dieses Problem kann mit dem sehr schmalen Prozessfenster für die Abscheidung einer stöchiometrischen dielektrischen Schicht beim reaktiven *Sputtern* erklärt werden, da ein solcher Prozess ständig Gefahr läuft, gänzlich zusammenzubrechen oder eine metallische Schicht zu erzeugen. Bei Stapel 2 und Stapel 3 kann damit im Laufe der weiteren Prozessierung also ein ähnlicher Effekt wie beim *Fencing* in Form von überschüssigem Aluminium und Al + Si-Clustern zum Tragen kommen, die es hier eigentlich zu vermeiden galt.

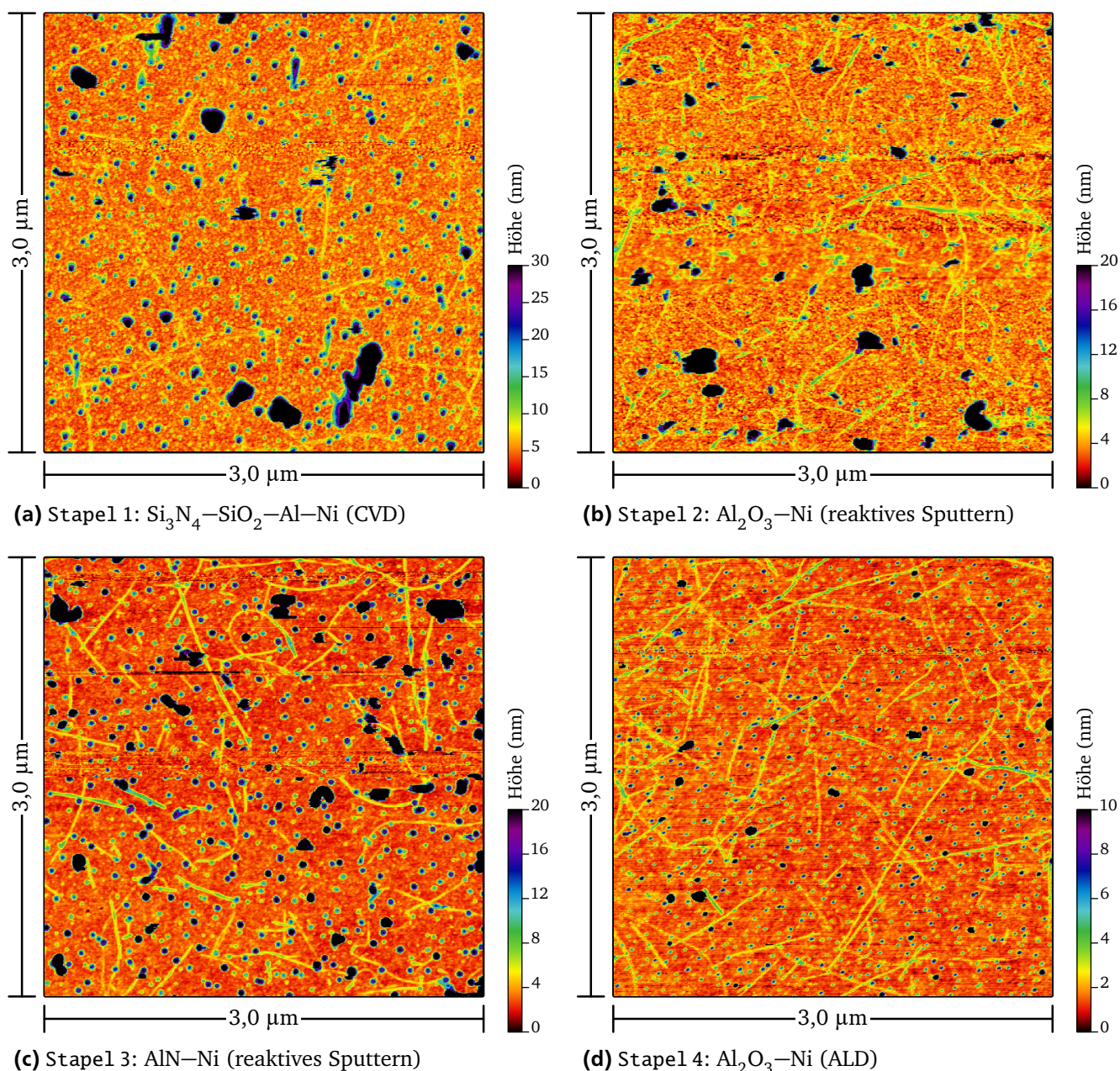


Abb. 65: AFM-Messung auf den verschiedenen Schichtstapeln: Die Cluster-Bildung und das CNT-Wachstum finden auf allen der hier untersuchten alternativen Schichtstapeln statt. Unter der Einschränkung, dass die gewählten Prozessparameter möglicherweise nicht für alle Schichtstapel optimal sind, wird für Stapel 4 hier das beste Ergebnis hinsichtlich Länge, Anzahl und Durchmesser der CNTs erzielt.

11 CNTFETs mit $\text{Al}_2\text{O}_3/\text{Ni}$ -Katalysator

11.1 Prozess zur Herstellung der FETs

Mit Hilfe der Untersuchungen in Kapitel 9 (Seite 93) konnte für das dielektrische Versagen der CNTFETs mit Aluminium/Nickel-Katalysator das sogenannte *Fencing* verantwortlich gemacht werden, das ein Artefakt der Strukturierung des Katalysators mittels *Lift off* ist. In Verbindung mit dem inhärenten Oxid-Verbrauch des Aluminium/Nickel-Katalysators ist *Fencing* der Grund für die äußerst geringe Ausbeute an funktionalen FETs. Deswegen wird hier der strukturierte Aluminium/Nickel-Katalysator durch einen unstrukturierten Aluminiumoxid/Nickel-Katalysator ersetzt. Der nun verwendete Prozess hat folgende Eckdaten, wobei die Parameter der Temperung und des CCVD-Schritts denen in Kapitel 10 (Seite 99) entsprechen:

- hoch Bor-dotierte Silizium-Substrate
- 100 nm thermisch gewachsenes Trocken-Oxid
- unstrukturierter Katalysator bestehend aus 5 nm Aluminiumoxid und 1 nm Nickel
- Sauerstoff-Plasmabehandlung des Katalysators
- Temperung in Argon bei $T_{\text{Temp}} = 870^\circ\text{C}$ für $t_{\text{Temp}} = 5 \text{ min}$
- CCVD mit Methan und Wasserstoff bei $T_{\text{CCVD}} = 870^\circ\text{C}$ für $t_{\text{CCVD}} = 10 \text{ min}$ (vgl. Abb. 66)
- *Source/Drain*-Kontakte mit 10 nm Palladium und 100 nm Nickel-Capping.
- Unterseite der Substrate als *Gate*-Kontakt (rückseitiges Oxid selektiv entfernt)

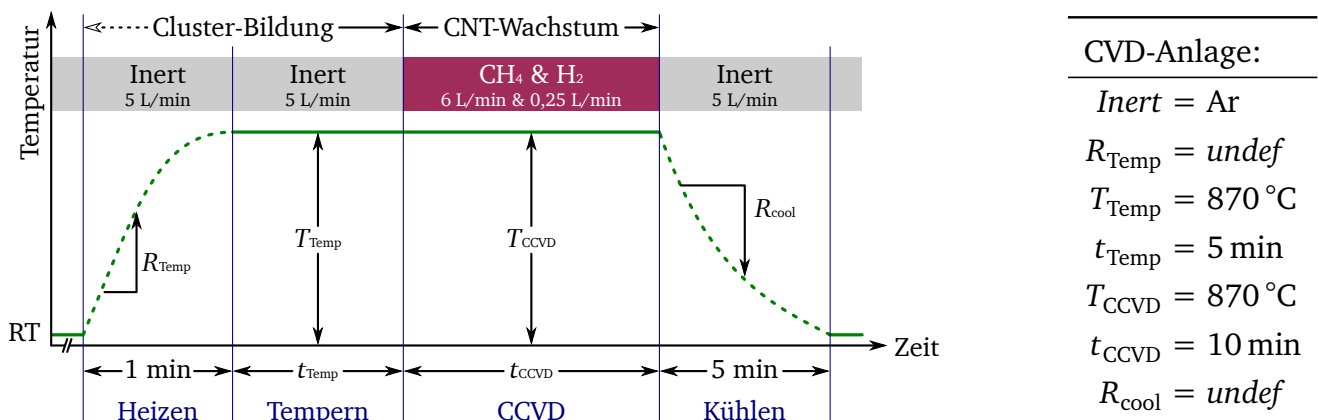


Abb. 66: Details der Prozessführung während der Temperung und des CCVD-Schritts für die Cluster-Bildung und das CNT-Wachstum zur Herstellung der CNTFETs mit Aluminiumoxid/Nickel-Katalysator.

Die Abscheidung des Aluminiumoxids erfolgt mit Hilfe der Atomlagenabscheidung (ALD) und 60 Zyklen des TMA/Wasser-Prozesses. Für die Abscheidung der Nickel-Schicht wird wie bisher das PVD-Verfahren der Elektronenstrahlverdampfung genutzt. Der ALD-Prozess wurde im Reinraum der TECHNISCHEN UNIVERSITÄT BERGAKADEMIE FREIBERG, und damit extern, mit der Anlage *Savannah S100* der Firma *UltraTech / Cambridge NanoTech* durchgeführt.

Obwohl durch die Verwendung des Aluminiumoxid/Nickel-Katalysators kein metallisches Aluminium vorhanden ist und deswegen auch kein Oxid-Verbrauch auftreten sollte, wird hinsichtlich der Prozessführung während der Temperung und während des CCVD-Schritts ein sehr konservativer Ansatz gewählt: Wie in Abb. 66 skizziert, werden sowohl die erhöhte Cluster-Bildungstemperatur T_{Temp} von 920 °C als auch die verlängerte CNT-Wachstumsdauer t_{CCVD} von 20 min auf die ursprünglichen Werte des Basisprozesses von 870 °C bzw. 10 min reduziert; einzig die Verwendung von Argon als Inertgas gegenüber der von Stickstoff wird beibehalten.

Es muss hier außerdem auf eine neuerliche Optimierung der Prozessführung für die Cluster-Bildung und das CNT-Wachstums analog zu Kapitel 7 verzichtet werden — die nun genutzten Parameter sind also für den verwendeten Katalysator nicht optimiert. Dass diese nicht optimierten Parameter ein CNT-Wachstum bewirken, das dem CNT-Wachstum unter Verwendung der für den alten Aluminium/Nickel-Katalysator optimierten Parameter entspricht, zeigen die beiden AFM-Messungen in Abb. 67: Abb. 67a zeigt Cluster und CNTs, die mit dem für den alten Aluminium/Nickel-Katalysator optimierten Prozess gebildet bzw. gewachsen wurden, während

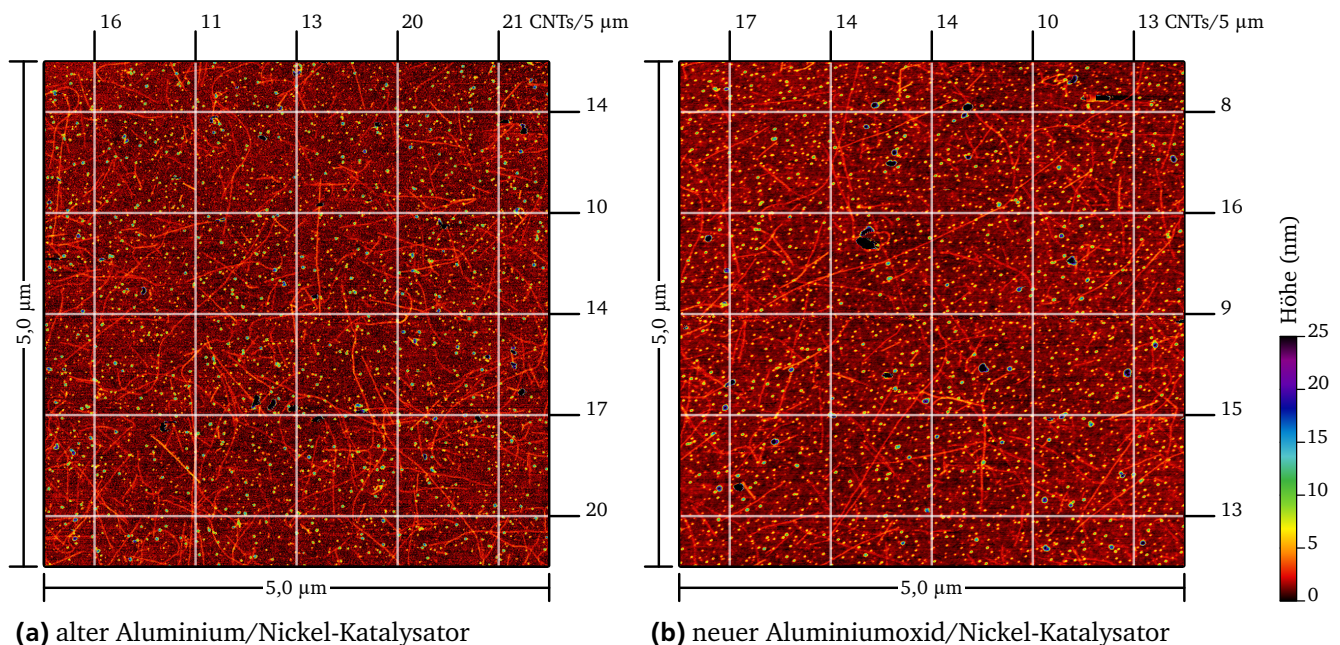


Abb. 67: Vergleich des CNT-Wachstums mit dem alten Aluminium/Nickel-Katalysator und dem neue Aluminiumoxid/Nickel-Katalysator: Hinsichtlich der (grob abgeschätzten) CNT-Dichten mit 3,1 CNTs/μm bzw. 2,6 CNTs/μm, der Länge der CNTs und des allgemeinen Gesamteindrucks wird mit dem neuen Katalysator ein dem alten Katalysator gleichwertiges Ergebnis erzielt.

Abb. 67b das Ergebnis für den nicht optimierten Prozess für den neuen Aluminiumoxid/Nickel-Katalysator zeigt. Für beide Messungen wurde die CNT-Dichte grob abgeschätzt, indem entlang der jeweils zehn eingezeichneten Linien die Schnittpunkte mit CNTs gezählt wurden. Dabei wird für den alten Katalysator eine CNT-Dichte von 3,1 CNTs/ μm erreicht, während die Dichte mit 2,6 CNTs/ μm für den neuen Katalysator leicht geringer ist. Aufgrund der Einschränkungen von AFM-Messungen, dünne CNTs bei nicht vernachlässigbarer Rauigkeit des Untergrunds nicht zuverlässig abzubilden, kann davon ausgegangen werden, dass für beide Messungen die tatsächliche CNT-Dichte höher ist. Deswegen und wegen der groben Abschätzung der CNT-Dichte, können die Ergebnisse für die beiden Katalysatoren als gleichwertig betrachtet werden; so sind für die Länge der CNTs und für den allgemeinen Gesamteindruck für beide Katalysatoren keine Unterschiede feststellbar.

11.2 Statistische Untersuchung der Ausbeute und des Schaltverhaltens

Zunächst sollen die Ausbeute und das Schaltverhalten von FETs, die mit dem neuen Aluminiumoxid/Nickel-Katalysator hergestellt wurden, statistisch betrachtet werden. Die in diesem Unterkapitel präsentierten Messungen wurde alle auf einem Wafer unter Verwendung des auf Seite 105 beschriebenen Prozesses gemacht. Die vermessenen FETs sind entlang der makroskopischen Orientierungshilfen des Testchip-Layouts (gelbe Markierungen in Abb. 35, Seite 56) angeordnet, wobei Tab. 8 die relevanten Daten dieser FETs zusammenfasst.

Die FETs vom Typ A haben eine nominale Kanallänge $L_{\text{Kanal|nom}}$ von $4\text{ }\mu\text{m}$ und eine nominale Kanalweite $W_{\text{Kanal|nom}}$ von $470\text{ }\mu\text{m}$. Bei den FETs vom Typ B handelt es sich um kleinere FETs mit einer Kanalweite von $20\text{ }\mu\text{m}$ mit der gleichen Kanallänge von $4\text{ }\mu\text{m}$. Die Fläche pro *Source*- oder *Drain*-Kontakt beträgt jeweils $50\,000\text{ }\mu\text{m}^2$ für die FETs vom Typ A bzw. $22\,500\text{ }\mu\text{m}^2$ für die FETs vom Typ B.

Die FETs vom Typ A sind jeweils in zwei parallelen Reihen senkrecht und parallel bzw. am und gegenüber des Flats angeordnet. Die insgesamt vier Reihen parallel zum Flat enthalten dabei jeweils 128 FETs, während die insgesamt vier Reihen senkrecht zum Flat 163 FETs enthalten. In der linken der beiden gegenüber des Flats gelegenen Reihen wurden die oberen 128 FETs für teilweise zerstörende Vorversuche verwendet, sodass hier in Summe nur 1036 der 1164 vorhandenen FETs vom Typ A vermessen werden konnten. Die FETs vom Typ B sind gleichmäßig in einzelnen Reihen von 166 Stück und einer Gesamtanzahl von 644 angeordnet, von denen alle vermessen wurden.

Die Messungen sind für die FETs vom Typ A in Abb. 69 und für die FETs vom Typ B in Abb. 70 zusammengestellt. Für jeden FET wurden der *Gate*-Strom und die Unterschwelkenennlinie und bei einer *Source/Drain*-Spannung V_{DS} von -500 mV und *Gate*-Spannungen V_{GS} von -4 bis $+4\text{ V}$ aufgenommen. Für jeden einzelnen FET wurden die folgenden Kennwerte ermittelt:

- der Strom I_{on} im niederohmschen bzw. *On*-Zustand als Median der Werte im Bereich von $V_{GS} = -4$ bis -3 V
- der Strom I_{off} im hochohmschen bzw. *Off*-Zustand als Median der Werte im Bereich von $V_{GS} = +3$ bis $+4\text{ V}$
- das *On/Off*-Verhältnis als Quotient der so bestimmten I_{on} und I_{off}
- der *Gate*-Leckstrom I_{Oxid} durch das Dielektrikum bei einer *Gate*-Spannung V_{GS} von $+4\text{ V}$

Auf Basis dieser Kennwerte werden die FETs, wie in Abb. 68 skizziert, klassifiziert, wobei die konkreten Grenzwerte graphisch aus Abb. 69 und Abb. 70 (rote Markierungen) extrahiert werden und zudem in Tab. 9 als absolute und auf die nominale Kanalweite bzw. auf die Kontakt-Fläche normierte Werte notiert sind.

- Ein FET ist funktional, wenn sein *Gate*-Leckstrom einen maximalen Grenzwert nicht überschreitet (erstes Kriterium) und wenn sein Strom im *On*-Zustand zusätzlich einen minimalen Grenzwert aufweist (zweites Kriterium).
- Ein FET hat eine halbleitende *Source/Drain*-Verbindung, schaltet also, wenn sein *On/Off*-Verhältnis mindestens einen Faktor 2 beträgt.
- Das Schaltverhalten wird in Größenordnungen des Stroms eingeteilt, wobei ein *On/Off*-Verhältnis von drei oder mehr Größenordnungen als „hoch“ bezeichnet wird.

Tab. 8: Nominale Kanallängen und -weiten und Kontakt-Flächen der FETs vom Typ A und der FETs vom Typ B.

	Anzahl	$L_{\text{Kanal nom}}$	$W_{\text{Kanal nom}}$	A_{pad}
Typ A	1036 Stück	4 μm	470 μm	50 000 μm^2
Typ B	664 Stück	4 μm	20 μm	22 500 μm^2

Tab. 9: Absolute und normierte Grenzwerte für die Klassifizierung der FETs vom Typ A und der FETs vom Typ B, die graphisch aus Abb. 69 und Abb. 70 extrahiert und über nominalen Kanalweiten bzw. die Kontakt-Flächen gemäß Tab. 8 berechnet wurden.

	absolut		normiert	
	$I_{\text{Oxid max}}$	$I_{\text{on min}}$	$I_{\text{Oxid max}}/(2 A_{\text{pad}})$	$I_{\text{on min}}/W_{\text{Kanal nom}}$
Typ A	2 nA	100 pA	20 fA/ μm^2	0,2 pA/ μm
Typ B	300 pA	100 pA	7 fA/ μm^2	5 pA/ μm

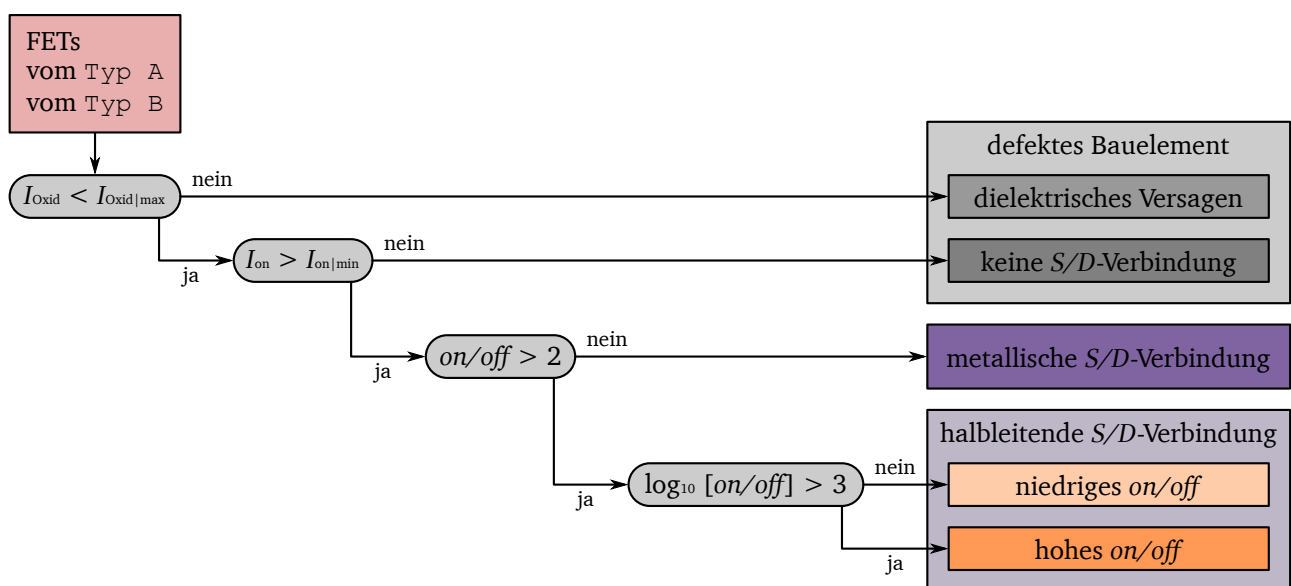
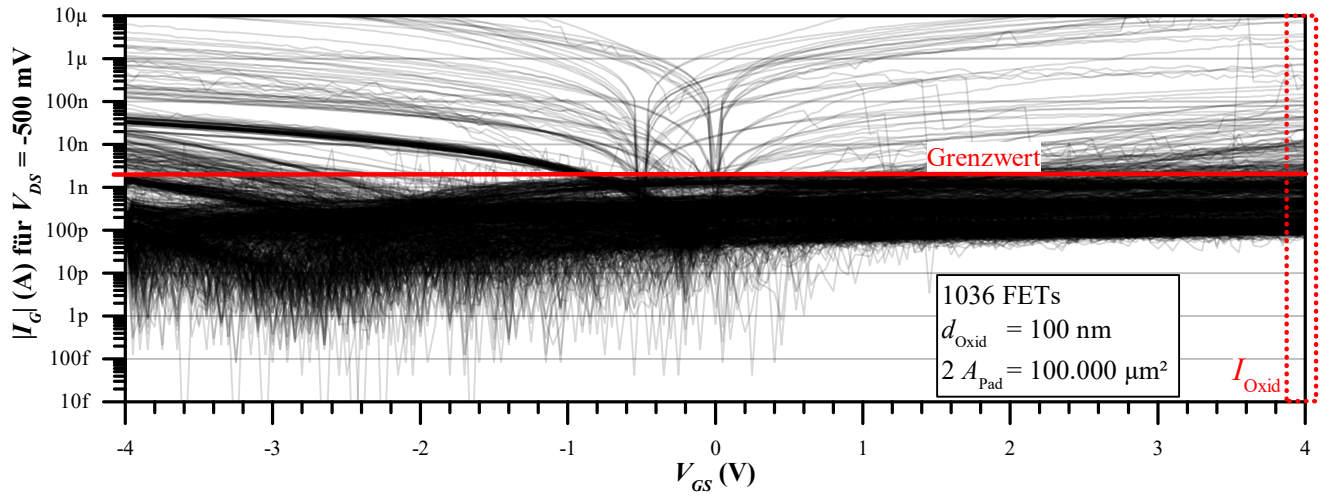
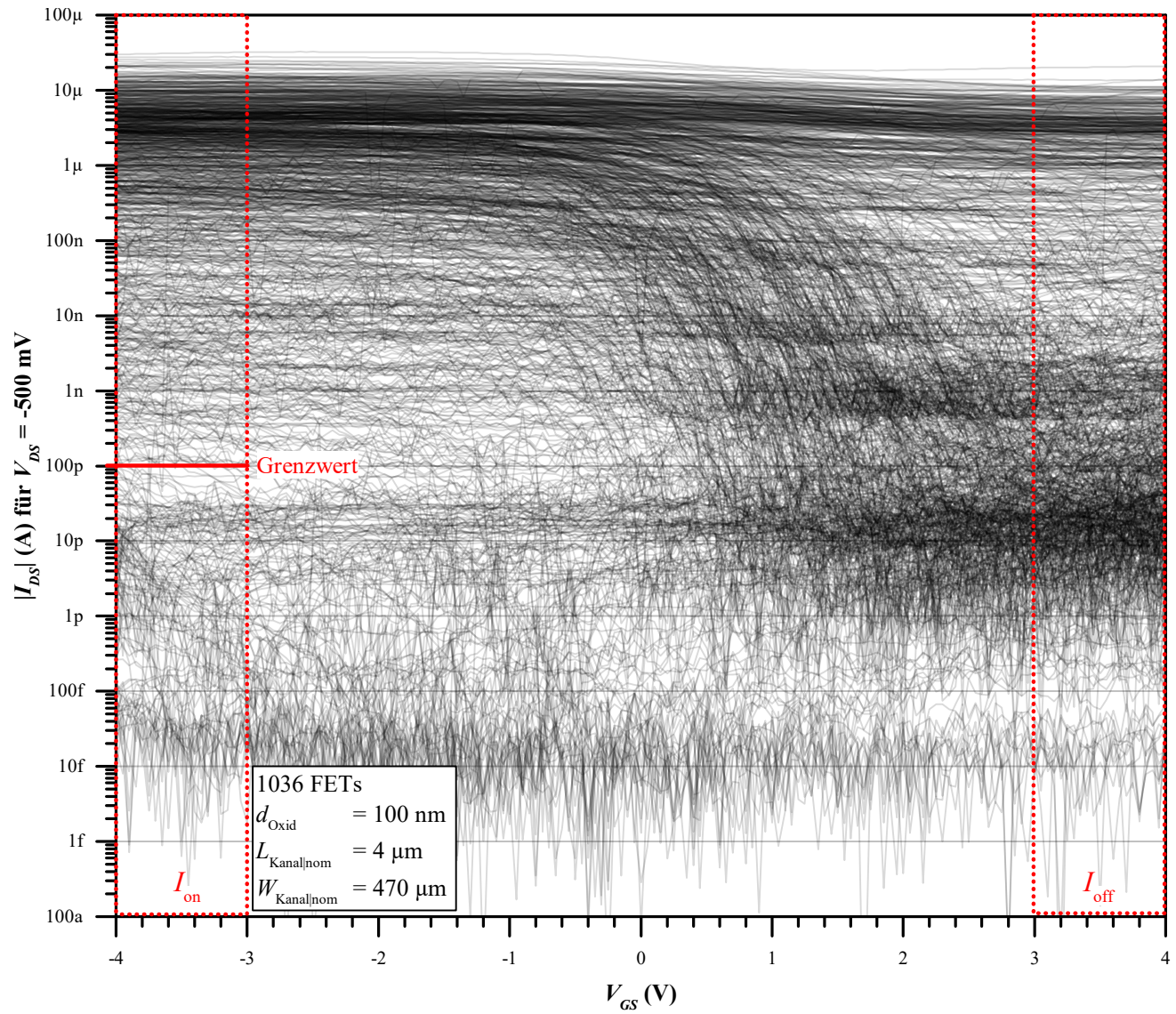


Abb. 68: Flussdiagramm zur Klassifizierung der FETs vom Typ A und der FETs vom Typ B basierend auf den aus Abb. 69 und Abb. 70 graphisch extrahierten Grenzwerten.

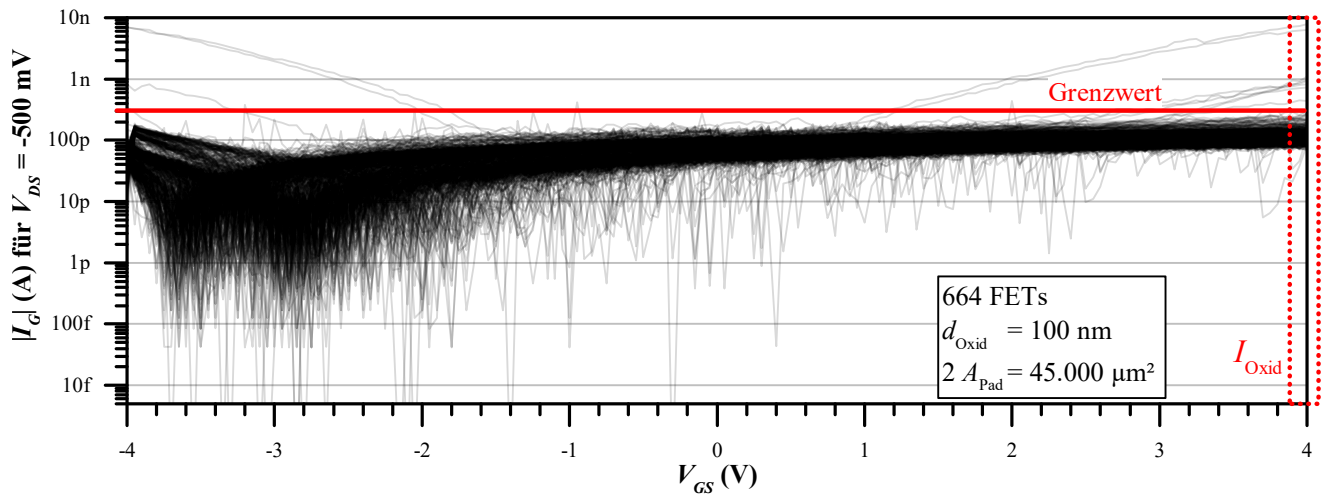


(a) Gate-Leckstrom: Ströme $I_{Oxid} \geq 2nA$ werden als dielektrisches Versagen des FET gewertet.

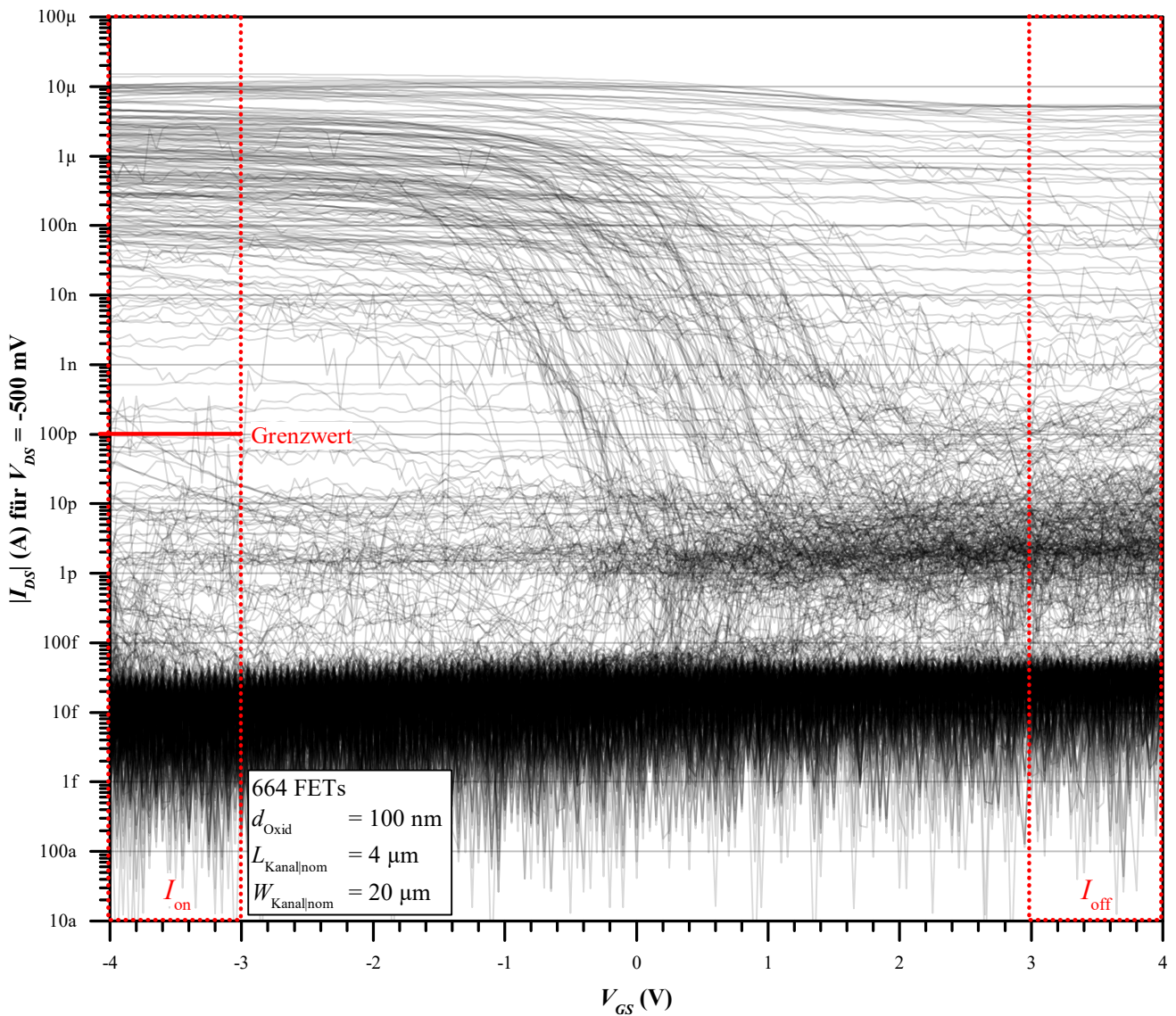


(b) Unterschwelkenennlinien: Ströme $I_{on} < 100pA$ werden als fehlende Source/Drain-Verbindung gewertet.

Abb. 69: Rohdaten der Messungen an 1036 FETs vom Typ A: I_G und I_{DS} bei $V_{DS} = -500 \text{ mV}$ für $V_{GS} = -4$ bis $+4 \text{ V}$.



(a) Gate-Leckstrom: Ströme $I_{\text{Oxid}} \geq 300 \text{ pA}$ werden als dielektrisches Versagen des FET gewertet.



(b) Unterschwelkenennlinien: Ströme $I_{\text{on}} < 100 \text{ pA}$ werden als fehlende Source/Drain-Verbindung gewertet.

Abb. 70: Rohdaten der Messungen an 664 FETs vom Typ B: I_G und I_{DS} bei $V_{DS} = -500 \text{ mV}$ für $V_{GS} = -4$ bis $+4 \text{ V}$.

Die Ergebnisse der Klassifizierung für die FETs vom Typ A sind in Abb. 71 graphisch dargestellt:

- Auf Basis von 1036 Bauelementen wird eine Ausbeute von fast 80 % erreicht, während ein Fünftel nicht funktional ist. Rund die Hälfte dieser Defekte werden durch das zweite Kriterium für funktionale Transistoren verursacht, d. h. dass insgesamt bei einem Zehntel aller Bauelemente in dieser Gruppe keine ausreichend leitfähige Verbindung zwischen *Source* und *Drain* zustande gekommen ist.
- Von den 808 funktionalen Bauelementen weist rund ein Fünftel eine metallische *Source/Drain*-Verbindung auf, d. h. sie schalten mit einem *On/Off*-Verhältnis von unter 2.
- Rund ein Viertel der 639 FETs mit halbleitender *Source/Drain*-Verbindung wird durch Bauelemente mit einem sehr geringen *On/Off*-Verhältnis von unter 10 gestellt. *On/Off*-Verhältnisse von einer, zwei, drei und vier Größenordnungen treten mit nahezu der gleichen Häufigkeit auf, während es für fünf und sechs Größenordnung nur wenige Bauelemente sind. Der Anteil an Transistoren mit hohem *On/Off* liegt damit bei 40 % der FETs mit halbleitender *Source/Drain*-Verbindung bzw. bei 24 % aller 1036 Bauelemente.
- Die 253 FETs mit hohem *On/Off*-Verhältnis weisen einen mittleren *On*-Strom von 733 nA und einen *Off*-Strom von 52 pA auf, um den sie mit 0,7 bzw. 0,9 Größenordnung streuen; im Mittel entspricht dies einem *On/Off* von $1,4 \times 10^4$.

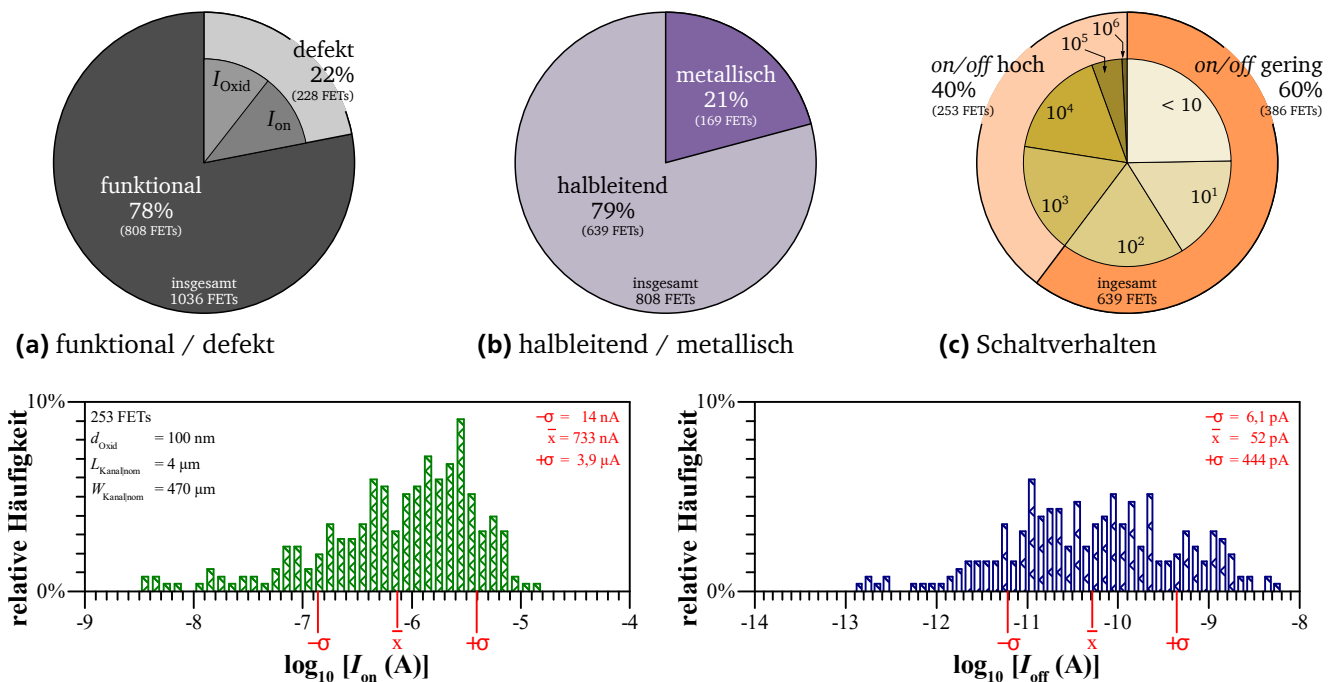


Abb. 71: Klassifizierung der 1036 FETs vom Typ A gemäß der Grenzwerte in Tab. 9 und mittels der Kennwerte, die aus den Messungen in Abb. 69 ermittelt wurden.

Die Ergebnisse der Klassifizierung für die FETs vom Typ B sind in Abb. 72 graphisch dargestellt:

- Auf Basis von 664 Bauelementen wird eine Ausbeute von 28 % erreicht; damit sind fast drei Viertel aller FETs als nicht funktional eingestuft. Nur ein Bruchteil dieser Defekte werden dabei durch das erste Kriterium für funktionale Transistoren verursacht, d. h. dass bei der Mehrheit aller Bauelemente das Dielektrikum zwar nicht durchgebrochen ist, aber keine ausreichend leitfähige Verbindung zwischen *Source* und *Drain* zustande gekommen ist.
- Von den 187 funktionalen Bauelementen weisen 13 % eine metallische *Source/Drain*-Verbindung auf, d. h. sie schalten mit einem *On/Off*-Verhältnis von weniger als 2.
- Knapp 80 % der 162 FETs mit halbleitender *Source/Drain*-Verbindung bzw. 19 % aller FETs vom Typ B sind Bauelemente mit hohem *On/Off*-Verhältnis von drei und mehr Größenordnungen. Die größte Gruppe stellen hier Bauelemente mit *On/Off*-Verhältnissen von vier oder fünf Größenordnungen. Rund ein Fünftel der halbleitenden FETs hat lediglich ein *On/Off* von zwei oder weniger Größenordnungen.
- Die 128 FETs mit hohem *On/Off*-Verhältnis weisen einen mittleren *On*-Strom von 450 nA und einen *Off*-Strom von 6,5 pA auf, um den sie mit 0,7 bzw. 1,1 Größenordnung streuen; im Mittel entspricht dies einem *On/Off* von $6,9 \times 10^4$.

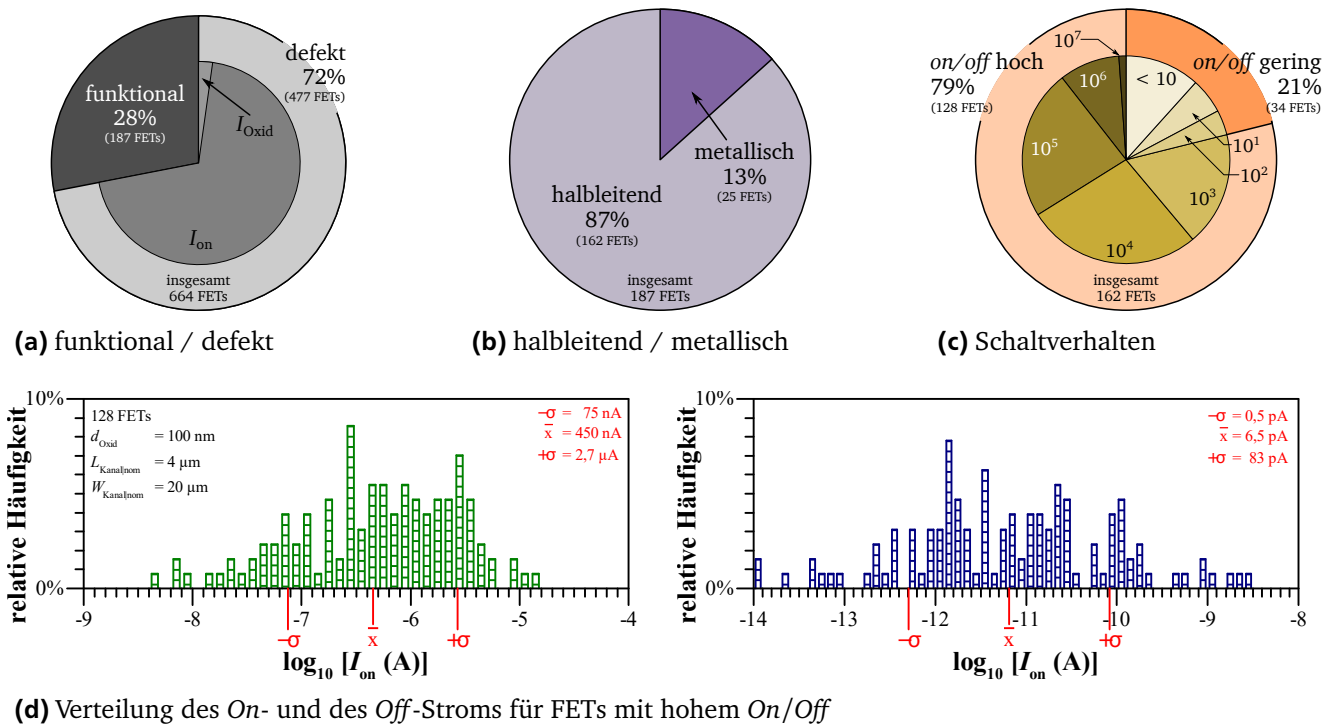


Abb. 72: Klassifizierung der 664 FETs vom Typ B gemäß der Grenzwerte in Tab. 9 und mittels der Kennwerte, die aus den Messungen in Abb. 70 ermittelt wurden.

Beim Vergleich der FETs vom Typ A und der FETs vom Typ B bzw. von Abb. 71a und Abb. 72a fallen zunächst die deutlich unterschiedlichen Ausbeuten auf: Werden bei den FETs vom Typ A 78 % funktionale Bauelemente erreicht, so stellt dieser Wert beinahe den prozentualen Anteil derjenigen Bauelemente dar, die bei den FETs vom Typ B als defekt klassifiziert wurden. Unter Beachtung des Flächenunterschieds der Kontakte von einem Faktor 2, wie in Tab. 8 notiert, wird dabei prozentual der gleiche Anteil an defekten Bauelemente durch das erste Kriterium, also ein dielektrisches Versagen des FETs, verursacht. Insbesondere zeigt auch Abb. 72a, dass die defekten Bauelemente in der Hauptsache durch das zweite Kriterium für funktionale Bauelemente bedingt sind, also dadurch dass keine leitfähige *Source/Drain*-Verbindung zustande gekommen ist. Da die nominale Kanallänge $L_{\text{Kanal|nom}}$ mit $4\text{ }\mu\text{m}$ für die FETs vom Typ A und FETs vom Typ B gleich ist, ist der relevante Unterschied zwischen den beiden Typen somit die nominale Kanalweite $W_{\text{Kanal|nom}}$, die für die FETs vom Typ B um den Faktor 23,5 geringer ist.

Die Wahrscheinlichkeit, dass eine leitende Verbindung zwischen *Source* und *Drain* zustande kommt (zunächst unabhängig davon, ob sie metallisch oder halbleitend ist, oder wie hoch ihr *On/Off*-Verhältnis ist), ist damit bei einem nominal $470\text{ }\mu\text{m}$ weiten Kanal höher als bei einem nur $20\text{ }\mu\text{m}$ weiten Kanal. Dieser Zusammenhang zwischen nominaler Kanalweite und Ausbeute kann mit einer zu geringen CNT-Dichte erklärt werden und ist zusätzlich in Abb. 73 stark vereinfacht skizziert: Ist die nominale Weite des Kanals groß, wie in Abb. 73a gezeigt, so ist die Wahrscheinlichkeit größer, dass sich CNTs so darin befinden, dass sie einen leitenden Pfad zwischen zwei Kontakten bilden, als wenn die nominale Kanalweite wie in Abb. 73b zum Beispiel nur halb so groß ist. Die Orientierung der CNTs auf eine einzige Richtung ist in Abb. 73 eine zeichnerische Vereinfachung des wirklichen CNT-Wachstums: tatsächlich sind die hier gewachsenen CNTs weder gerade noch auf eine Richtung beschränkt (vgl. etwa Abb. 67b). Der skizzierte Zusammenhang zwischen nominaler Kanalweite und Ausbeute bleibt dennoch gültig, da eine zufällige Orientierung der CNTs die Ausbeute für kleinere Kanalweiten zusätz-

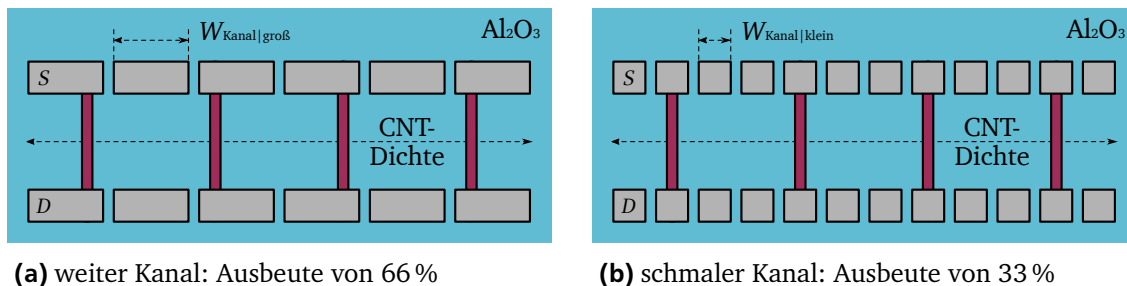


Abb. 73: Vereinfachte schematische Darstellung des Zusammenhangs zwischen der Ausbeute an funktionalen Bauelementen und nominaler Kanalweite: Bei gleicher CNT-Dichte wird durch weite Kanäle eine höhere prozentuale Ausbeute erreicht als für schmale Kanäle. Die Orientierung der CNTs auf eine einzige Richtung ist hierbei eine zeichnerische Vereinfachung; tatsächlich sind die hier gewachsenen CNTs weder gerade noch auf eine Richtung beschränkt, vgl. etwa Abb. 67b (Seite 106).

lich reduziert, indem beispielsweise eine leitfähige Verbindung zwischen *Source* und *Drain* von verschiedenen FETs zustande kommt; analoge Überlegungen gelten für gekrümmte CNTs.

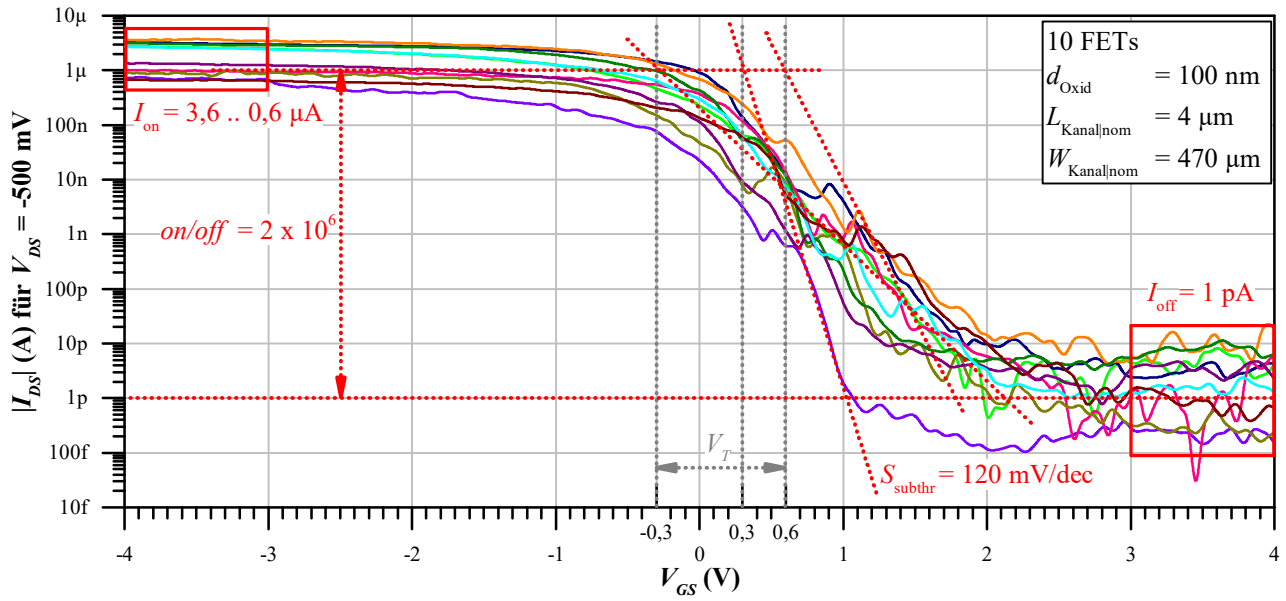
In der Gruppe der funktionalen Bauelemente zeigen Abb. 71b und Abb. 72b für die FETs vom Typ A und die FETs vom Typ B nahezu die gleiche anteilige Zusammensetzung von metallischen und halbleitenden *Source/Drain*-Verbindungen. Dieser Zusammenhang kann damit erklärt werden, dass hier nur zustande gekommene leitfähige Verbindungen zwischen *Source* und *Drain* betrachtet werden und diese sich aus einer gemeinsamen CNT-Population bilden. Ob eine *Source/Drain*-Verbindung also metallisches oder halbleitendes Verhalten zeigt, ist damit von der nominalen Kanalweite unabhängig. Dass sich die tatsächlichen Anteile an metallischen Verbindungen für die beiden Typen unterscheiden, kann durch zwei Umstände erklärt werden: zunächst ist der Umfang der Stichprobe für die FETs vom Typ B aufgrund der geringen Ausbeute deutlich geringer, sodass das Ergebnis einen höheren statistischen Fehler aufweisen könnte und der Wert damit auch größer als 13 % sein kann; außerdem kann es für die FETs vom Typ A aufgrund der größeren nominalen Kanalweite in einigen Fällen zu einem nicht vollständigen *Lift off* gekommen sein, sodass metallische Filamente zwischen den Kontakten zurückgeblieben sind, die dann den Anteil metallischer Verbindungen parasitär erhöhen; somit kann der tatsächliche Anteil an metallischen Verbindungen geringer als 21 % sein.

Hinsichtlich des Schaltverhaltens der FETs vom Typ A und FETs vom Typ B zeigen Abb. 71c und Abb. 72c, dass der Anteil an Bauelementen mit hohem *On/Off*-Verhältnis für die FETs vom Typ B deutlich höher ist. Da das Schaltverhalten von derjenigen *Source/Drain*-Verbindung bestimmt wird, die am schlechtesten schaltet, nimmt die Wahrscheinlichkeit ein geringes *On/Off*-Verhältnis für ein Bauelement zu erhalten, mit der Zahl der Verbindungen in diesem Bauelement zu. Und da die Wahrscheinlichkeit eine hohe Anzahl von Verbindungen in einem Bauelement zu haben, mit der nominalen Kanalweite zunimmt, nimmt auch die Wahrscheinlichkeit für ein geringes *On/Off*-Verhältnis zu.

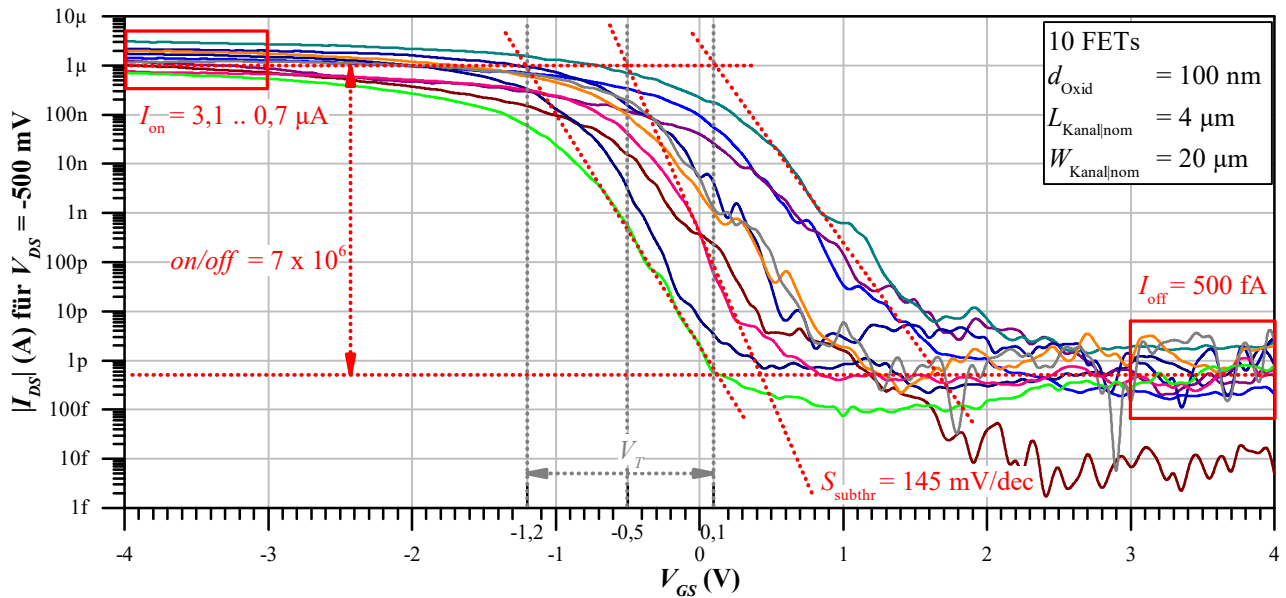
11.3 Untersuchungen von ausgewählten FETs

Auf Basis der statistischen Untersuchungen wurden jeweils 10 FETs vom Typ A und FETs vom Typ B ausgewählt, deren Unterswellenkennlinien in Abb. 74 dargestellt sind. Die Kennlinien wurden bei $V_{DS} = -500 \text{ mV}$ für V_{GS} von -4 bis $+4 \text{ V}$ aufgenommen, während die nominale Kanalweite $W_{\text{Kanal|nom}}$ $470 \text{ }\mu\text{m}$ bzw. $20 \text{ }\mu\text{m}$ und nominale Kanallänge $L_{\text{Kanal|nom}}$ wie gehabt $4 \text{ }\mu\text{m}$ beträgt.

Die 10 FETs vom Typ A in Abb. 74a zeigen dabei *On*-Ströme I_{on} im Bereich von $3,6$ bis $0,6 \text{ }\mu\text{A}$ bzw. *On*-Widerstände R_{on} von 140 bis $830 \text{ k}\Omega$, während die *Off*-Ströme I_{off} im Mittel bei 1 pA liegen, was einem mittleren *Off*-Widerstand R_{off} von $500 \text{ G}\Omega$ entspricht. Damit kann ein mittleres



(a) Messungen an 10 ausgewählten FETs vom Typ A



(b) Messungen an 10 ausgewählten FETs vom Typ B

Abb. 74: Unterschwellenkennlinien, gemessen bei $V_{DS} = -500$ mV für $V_{GS} = -4$ bis $+4$ V.

On/Off-Verhältnis von 2×10^6 berechnet werden. Die maximal erreichte Unterschwellensteigung S_{subthr} beträgt 120 mV/dec, während im Mittel Werte von 220 mV/dec erreicht werden, was gute bis sehr gute Werte im Vergleich zu den bisherigen Ergebnissen dieser und der Vorgängerarbeit darstellt. Für den FET mit der höchsten Unterschwellensteigung kann eine Schwellenspannung V_T von $+0,3$ V ermittelt werden, um die die anderen Transistoren im Bereich von $-0,3$ bis $+0,6$ V streuen. Die 10 FETs vom Typ B in Abb. 74b zeigen On-Ströme I_{on} im Bereich von 3,1 bis 0,7 μA bzw. On-Widerstände R_{on} von 160 bis 715 k Ω , während die Off-Ströme I_{off} im Mittel bei 500 fA liegen, was einem mittleren Off-Widerstand R_{off} von 1000 G Ω entspricht. Damit kann ein mittleres On/Off-Verhältnis von 7×10^6 berechnet werden. Die maximal erreichte

Unterschwellensteigung S_{subthr} beträgt 145 mV/dec, während im Mittel Werte von 250 mV/dec erreicht werden, was immer noch gute aber gegenüber den FETs vom Typ A leicht reduzierte Werte darstellt. Die FETs vom Typ B weisen eine stärker streuende Schwellspannung V_T im Bereich von $-1,2$ bis $+0,1$ V, wobei für den FET mit der höchsten Unterschwellensteigung eine Schwellspannung von $-0,5$ V bestimmt werden kann.

11.4 Elektrische Zerstörung metallischer CNTs

Für die meisten CVD-Prozesse, die für das Wachstum von CNTs genutzt werden können, beträgt der Anteil an halbleitenden CNTs zwei Drittel gegenüber einem Drittel an CNTs mit metallischem Verhalten. Die hier in Kapitel 11.2 (Seite 108) gemachten statistischen Untersuchungen zeigen, dass rund 15 bis 20 % der Bauelemente metallisches Verhalten zeigen. Aufgrund der hohen nominalen Kanalweiten kann für diese Bauelemente und für Bauelemente mit nur geringem *On/Off*-Verhältnis vermutet werden, dass das schlechte Schaltverhalten von Kurzschlüssen durch metallische CNTs verursacht wird. Daher ist es wünschenswert, die nicht-halbleitenden CNTs in bereits vollständig prozessierten Bauelementen nachträglich auszusondern zu können. Eine solche Aussonderung von metallischen CNTs ist für diese FETs unumgänglich, wenn ein gutes Schaltverhalten erreicht werden soll. Da aber mit vertretbarem Aufwand weder eine individuelle Identifizierung der metallischen CNTs noch deren mechanische Zerstörung möglich ist, ist ein Verfahren gesucht, das auf elektrischem Weg selektiv metallische CNTs zerstört.

Die hier beschriebene und speziell für diese Arbeit entworfene Anordnung zur elektrischen Zerstörung von metallischen CNTs („Impuls-Generator“ bzw. „Impuls-Überlastung“) ist so ausgelegt, dass sie in das in Kapitel 4.3 (Seite 32) vorgestellte System zur elektrischen Charakterisierung integriert werden kann. Es ist damit möglich, hergestellte Bauelemente elektrisch zu charakterisieren, unerwünschte metallische CNTs mit Hilfe des Impuls-Generators zu zerstören, das Bauelement anschließend erneut zu charakterisieren und die Impuls-Überlastung bei Bedarf zu wiederholen. Die Überlastung erfolgt durch einen kurzen Strom-Impuls, der zu einer massiven lokalen Erhitzung und Zerstörung elektrisch leitfähiger CNTs durch Joule'sche Eigenheizung führt [97]. In Atmosphäre liegen die dafür nötigen Temperaturen im Bereich von 500 bis 700 °C [98]. Andere Arbeiten berichten, dass es bei wiederholter Anwendung von Pulsen möglich sei, die metallischen Lagen einer MWNT selektiv zu zerstören und nur die halbleitenden Lagen zu erhalten, sodass durch „Abschälen“ der metallischen Lagen die Transformation einer metallischen MWNT in eine halbleitende SWNT möglich werde [99, 100].

Ein detaillierter Schaltplan des Impuls-Generators ist samt Platinenlayout in den Anhang auf Seite 139 gestellt, während die wesentlichen fünf Blöcke hier in Abb. 75 skizziert sind. Im Einzelnen läuft der Vorgang dabei wie folgt ab: Eine Versorgungsspannung V_{Supply} wird zwi-

schen den *Source/Drain*-Kontakten des Bauelements mit den zu zerstörenden metallische CNTs angelegt („DUT“, *device under test*); dann wird der zunächst noch offene Stromkreis für eine definierte Dauer t_{Puls} geschlossen, wobei ein Überlastungs-Strom I_{DUT} durch das Bauelement samt seiner metallischen CNTs fließt. Da ein Teil der Versorgungsspannung V_{Supply} auch über andere Elemente der Schaltung abfallen kann, wird die tatsächlich am Bauelement anliegende Spannung V_{DUT} separat gemessen. Während des gesamten Vorgangs ist an das *Gate* des Bauelements eine geeignete Spannung V_{GS} angelegt, die die halbleitenden CNTs in ihrem *Off*-Zustand hält. Damit ist die Leitfähigkeit der halbleitenden CNTs gegenüber der Leitfähigkeit der metallischen CNTs soweit herabgesetzt, dass der Überlastungs-Strom I_{DUT} effektiv nur durch die metallischen CNTs fließt und somit auch nur diese thermisch zerstört.

Der Strom-Impuls erzeugt allerdings aufgrund parasitärer Induktivitäten und Kapazitäten und fehlender Impedanzanpassung der Zuleitungen störende Messartefakte. Wegen dieser Artefakte ist eine Messung des zeitlichen Verlaufs des Strom-Impulses I_{DUT} und der Spannung V_{DUT} nicht möglich. Daher muss hier unter Vernachlässigung dynamischer Effekte stark vereinfachend mit der nominalen *Peak*-Leistung P_{Puls} anstelle der effektiven Energie E_{Puls} argumentiert werden. Dafür wird zunächst davon ausgegangen, dass die gesamte Versorgungsspannung V_{Supply} über das Bauelement abfällt, d. h. es gilt $V_{\text{DUT}} = V_{\text{Supply}}$. Bei Kenntnis des initialen Widerstands des Bauelements R_{DUT} ist die nominale *Peak*-Leistung P_{Puls} des Überlastungs-Impulses dann nach Gl. 49 bestimmbar. Der initiale Widerstand R_{DUT} des Bauelements muss dafür vor der Impuls-Überlastung separat über eine Unterschwellenkennlinie im *Off*-Zustand des Transistors bestimmt werden.

$$P_{\text{Puls}} = \frac{V_{\text{DUT}}^2}{R_{\text{DUT}}} \quad \text{wobei} \quad R_{\text{DUT}} = \frac{V_{\text{DS}}}{I_{\text{off}}} \quad (49)$$

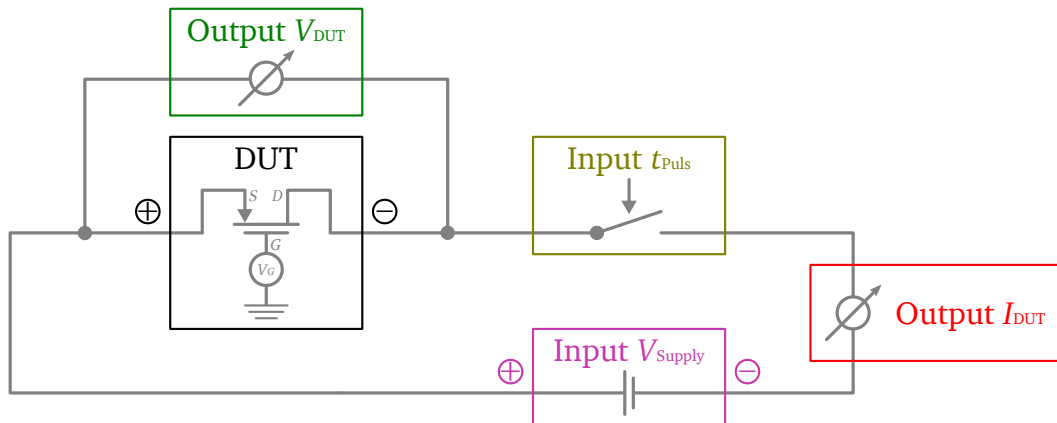


Abb. 75: Aufbau des Impuls-Generators zur elektrischen Zerstörung metallischer CNTs: Sobald der Stromkreis für die Dauer t_{Puls} geschlossen wird, fließt aufgrund der angelegten Spannung V_{Supply} ein Strom I_{DUT} durch die metallischen CNTs des Bauelements (DUT); die halbleitenden CNTs des Bauelements werden durch eine geeignete Spannung V_{GS} im *Off*-Zustand gehalten und somit geschützt.

Abb. 76 zeigt die Unterschwellenkennlinien des Transistors FET 3. Es handelt sich um einen Fingertransistor mit einer nominalen Kanallänge $L_{\text{Kanal|nom}}$ von $2\text{ }\mu\text{m}$ und mit 32 Fingerpaaren, was einer nominalen Kanalweite $W_{\text{Kanal|nom}}$ von $5760\text{ }\mu\text{m}$ entspricht. Für alle Kennlinien wurde der *Source/Drain*-Strom I_{DS} bei der jeweils gleichen *Source/Drain*-Spannung V_{DS} von -500 mV für *Gate*-Spannungen V_{GS} von -4 bis $+4\text{ V}$ gemessen.

Die blaue Kurve zeigt dabei das initiale Verhalten des FET. In diesem Zustand weist der Transistor einen *On*- und *Off*-Strom von $26,4\text{ mA}$ bzw. $25,9\text{ mA}$ auf, was einem *de facto* nicht vorhandenem Schaltverhalten mit einem *On/Off*-Verhältnis von einem Faktor $1,02$ entspricht. Nach dieser ersten Messung, aus der ein Widerstand R_{DUT} von $19\text{ }\Omega$ für den Transistor bestimmt werden kann, wird der Impuls-Generator genutzt, um die metallischen CNTs elektrisch zu zerstören und das Schaltverhalten des Transistors zu verbessern. Dabei wird der Transistor zunächst in seinen *Off*-Zustand geschaltet und dann ein Überlastungs-Impuls mit $V_{\text{DUT}} = 3\text{ V}$ und einer nominalen Dauer t_{Puls} von $200\text{ }\mu\text{s}$ angelegt, was einer nominalen *Peak*-Leistung P_{Puls} von 470 mW entspricht. Die zweite Messung (grüne Kurve) zeigt nun reduzierte Ströme im *On*- und *Off*-Zustand von 2 mA bzw. $600\text{ }\mu\text{A}$ und ein höheres, aber immer noch sehr geringes *On/Off*-Verhältnis von $3,33$. Es lässt sich nun allerdings ein eindeutig unipolares p-MOS-Verhalten des Transistors fest-

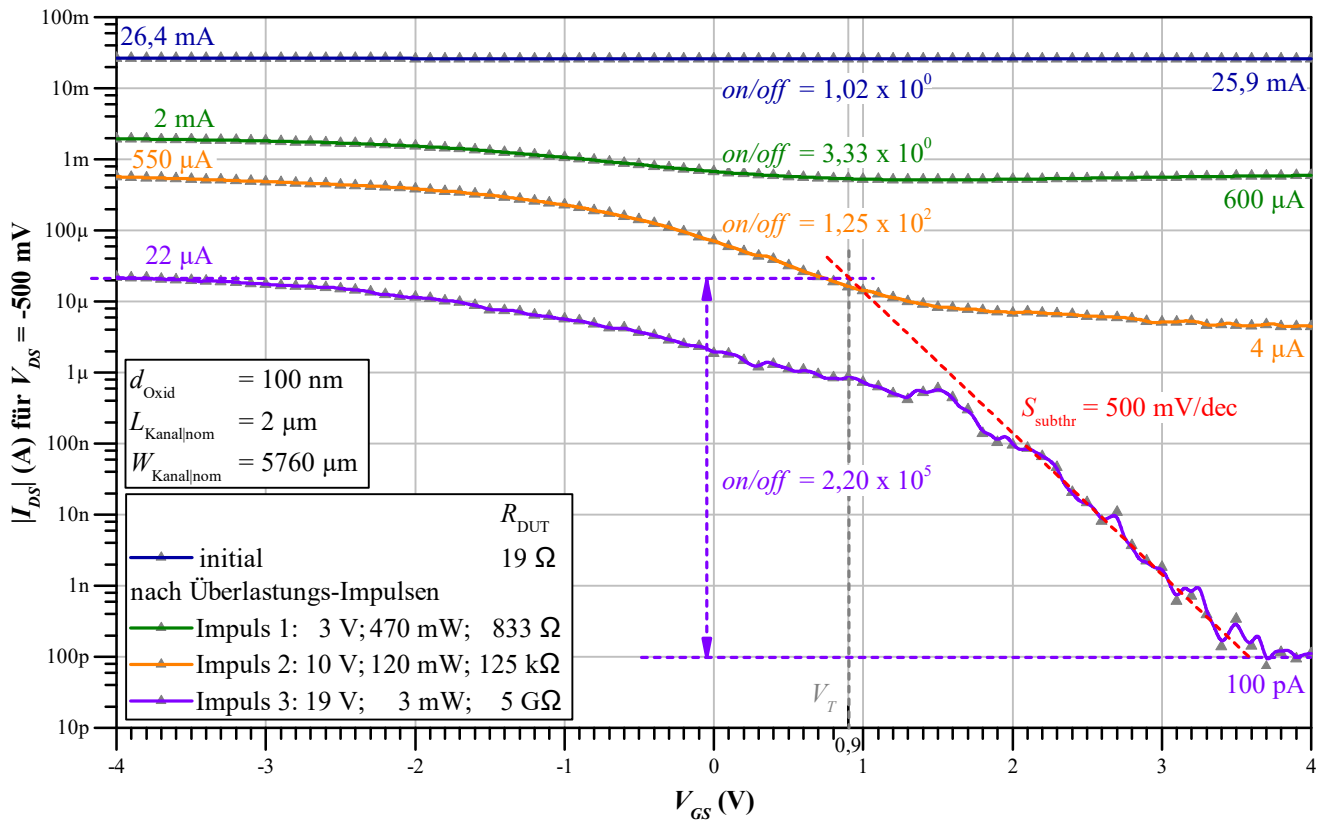


Abb. 76: Elektrische Zerstörung von metallischen CNTs in FET 3: Bei $V_{\text{DS}} = -500\text{ mV}$ gemessene Unterschwellenkennlinien des Transistors mit initialem Verhalten und nach sukzessiver Anwendung von drei Überlastungs-Impulsen. Die Impulse haben jeweils nominale Leistungen im Bereich von 470 bis 3 mW ; während der Impuls-Überlastung wird der Transistor in den *Off*-Zustand geschaltet.

stellen. Nach einem weiteren Impuls mit 10 V bzw. 120 mW zeigt die dritte Messung (orange Kurve) nochmals reduzierte Ströme; nun aber mit einem *On/Off*-Verhältnis von rund zwei Größenordnungen. Ein dritter und letzter Impuls mit 19 V bzw. 3 mW erzeugt einen Transistor, der sich entsprechend der violetten Kurve verhält. In seinem finalen Zustand hat der Transistor nun einen *On*- und *Off*-Strom von 22 μA bzw. 100 pA, was einem *On*- und *Off*-Widerstand von 23 k Ω bzw. 9 G Ω entspricht. Damit kann ein Schaltverhalten mit einem *On/Off*-Verhältnis von rund fünf Größenordnungen berechnet werden. Die Unterschwellensteigung S_{subthr} liegt bei 500 mV/dec, woraus die Schwellspannung V_T zu $-0,3\text{ V}$ bestimmt werden kann.

11.5 Bewertung der Ergebnisse

Basierend auf der in diesem Kapitel durchgeführten statistischen Auswertung der 1036 FETs vom Typ A und der 664 FETs vom Typ B und der Auswertung von verschiedenen ausgewählten FETs aus beiden Gruppen soll die Herstellung von CNTFETs auf Basis von *in situ* gewachsenen CNTs nun bewertet werden.

CNT-Wachstum unter Verwendung des neuen Katalysators

Der Umstieg vom alten Aluminium/Nickel-Katalysator auf den neuen Aluminiumoxid/Nickel-Katalysator hat für das CNT-Wachstum keine schwerwiegenden negativen Konsequenzen. Dies zeigen die AFM-Messung in Abb. 67b (Seite 106) und die elektrischen Messungen in diesem Kapitel. Dennoch hat Kapitel 7 (Seite 67) für den alten Katalysator gezeigt, dass es möglich ist, die Cluster-Bildung und das CNT-Wachstum zu optimieren, indem die Parameter der Temperung und des CCVD-Schritt untersucht und angepasst werden. Eine Untersuchung analog Kapitel 7 sollte also auch für den neuen Katalysator durchgeführt werden.

Testchip-Layout: Ausbeute und Schaltverhalten der CNTFETs

Der statistisch festgestellte Zusammenhang zwischen nominaler Kanalweite und Ausbeute an funktionalen Bauelementen bzw. deren Schaltverhalten lässt zwei gegenläufige Entwicklungen erkennen: Werden für die FETs vom Typ A, die eine nominale Kanalweite von 470 μm aufweisen, rund 80 % an funktionalen Bauelementen erreicht, so sind es für die FETs vom Typ B mit einer Kanalweite von 20 μm nur rund 30 %. Dagegen erreichen für die FETs vom Typ B rund 80 % ein hohes *On/Off*-Verhältnis, während es bei den FETs vom Typ A nur 40 % sind.

Außerdem kommt hinzu, dass durch Erhöhung der nominalen Kanalweiten gegenüber dem Layout der Vorgängerarbeit nur eine generell unterproportionale Erhöhung des *On*-Stroms erreicht wurde. Eine elektrische Parallelisierung von mehreren CNTs in einem Transistor findet also nicht in dem erhofften Umfang statt.

Einsatz des Impuls-Generators

Die an FET 3 gemachten Messungen zeigen, dass es möglich ist einen Transistor, der einen hohen *On*-Strom, aber kein Schaltverhalten zeigt, so zu bearbeiten, dass ein hohes *On/Off*-Verhältnis erzeugt wird, während gleichzeitig der *On*-Strom nicht proportional reduziert wird. Die Unterschwellensteigung des resultierenden Transistors erreicht allerdings nur schlechte Werte. CNTFETs, die ein hinsichtlich ihres *On/Off*-Verhältnis *und* ihrer Unterschwellensteigung gutes Schaltverhalten zeigen, sind unter Einsatz des Impuls-Generators in seiner derzeitigen Form also nicht zu erreichen. Zudem ist die Impuls-Überlastung trotz der hier komfortabel gelösten Integration in die Messtechnik ein zeitaufwendiger Prozess, da nicht alle störenden, d. h. metallischen, CNTs mit nur einem Impuls zerstört werden können, sondern die Überlastung in mehreren Stufen durchgeführt werden muss. Für Herstellungstechnologien, die über den Labormaßstab hinausgehen, muss die Impuls-Überlastung also optimiert und automatisiert werden. Der dafür wichtigste Punkt ist eine zeitaufgelöste Messung des Strom-Impulses der Überlastung, um die aufgewendete Energie präzise bestimmen zu können.

12 Zusammenfassung

Aufbauend auf der Vorgängerarbeit [6] konnte in dieser Arbeit eine Prozesstechnologie samt Testchip-Layout zur Herstellung von Kohlenstoffnanoröhren-Feldeffekt-Transistoren (CNTFETs) entwickelt werden. Das zentrale Element ist dabei die katalytisch-chemische Gasphasenabscheidung (CCVD) zum *in situ*-Wachstum von Kohlenstoffnanoröhren (CNTs) unter Verwendung von Methan als Kohlenstoff-Quelle. Der Prozess nutzt drei Layout-Ebenen, die die Strukturierung des Katalysators in Form von Inseln ermöglichen, auf die *Source/Drain*-Kontakte justiert werden. Die Strukturierung der Katalysator-Inseln und der *Source/Drain*-Kontakte erfolgt mittels der *Lift off*-Technik. Anstelle individueller *Topgates* wird ein *Backgate*-Ansatz verfolgt, wodurch auch an fertig prozessierten Bauelementen noch topologische Untersuchungen mittels Rasterkraftmikroskopie (AFM) durchführbar sind. Bezüglich des verwendeten Katalysators ist der Prozess flexibel anpassbar.

Zunächst wurde mit einem Aluminium/Nickel-Katalysator gearbeitet und mittels AFM die Cluster-Bildung und das CNT-Wachstum untersucht. Ziel war dabei, ausgehend vom Basisprozess Parameter für die Temperung und den CCVD-Schritt zu finden, die zum Wachstum einer hohen Anzahl an CNTs mit möglichst geringem Durchmesser führen. Für eine optimale Cluster-Bildung wurden hohe Heizraten, eine Temperatur von 920 °C und die Verwendung von Argon als Inertgas ermittelt. Außerdem hat sich eine Sauerstoff-Plasmabehandlung des Katalysators vor der Temperung als vorteilhaft herausgestellt. Für das CNT-Wachstums wurde festgestellt, dass bei Verlängerung der Wachstumsdauer mehr CNTs gewachsen werden, deren Länge davon aber unbeeinflusst bleibt.

Für CNTFETs, die unter Verwendung der optimierten Parameter für Temperung und CCVD-Schritt und des Aluminium/Nickel-Katalysators gefertigt wurden, wurde nur eine sehr geringe Ausbeute an funktionalen Bauelementen erreicht, weswegen diese Bauelemente nur selten elektrisch charakterisiert werden konnten. Als Ursache für die geringe Ausbeute wurde eine gravierende Schädigung des *Gate*-Oxids festgestellt, für die wiederum überschüssiges Aluminium an den Katalysator-Inseln verantwortlich gemacht werden kann. Das überschüssige Aluminium wurde dabei als Artefakt der Strukturierung mittels *Lift off* erkannt.

Um die Ausbeute an funktionalen Bauelementen zu erhöhen, wurden in der Folge mehrere Maßnahmen ergriffen: Zunächst wurde die Strukturierung des Katalysators mittels *Lift off* aufgegeben und durch einen vollflächigen Katalysator ohne Inseln ersetzt. Außerdem wurde das metallische Aluminium des Katalysators gegen ein dielektrisches Aluminiumoxid ausgetauscht,

das mittels Atomlagenabscheidung (ALD) abgeschieden wurde. Schließlich wurden die meisten Optimierungen der Parameter von Temperatur und CCVD-Schritt zurückgenommen und im Folgenden mit den Parametern des Basisprozesses gearbeitet. Hierbei wurden die gegenüber dem Basisprozess erhöhte Cluster-Bildungs-Temperatur und die verlängerte CNT-Wachstumsdauer nicht weiter verwendet.

Für CNTFETs, die unter Verwendung des unstrukturierten Aluminiumoxid/Nickel-Katalysators hergestellt wurden, konnte auf Basis statistischer Untersuchungen eine Korrelation zwischen der nominalen Kanalweite und der Ausbeute an funktionalen Bauelemente und deren Schaltverhalten festgestellt werden: Je höher die Kanalweite der Transistoren ist, desto höher ist die Ausbeute an funktionalen Bauelementen, desto schlechter ist allerdings auch deren Schaltverhalten. Unabhängig von der nominalen Kanalweite haben rund ein Fünftel der hergestellten Transistoren eine metallische *Source/Drain*-Verbindung. Mittels eines eigens für diese Arbeit entworfenen Impuls-Generators konnte exemplarisch an einem Bauelement mit metallischer *Source/Drain*-Verbindung gezeigt werden, dass das zunächst rein metallische Verhalten in ein Schaltverhalten mit einem *On/Off*-Verhältnis von fünf Größenordnungen gesteigert werden kann.

Anhang

Herleitung des CNT-Durchmessers

Der Durchmesser d_{SWNT} einer CNT ist über deren Umfang U_{SWNT} gegeben, wobei dieser gleich der Länge $|\vec{C}|$ des Chiralitätsvektors ist. Es ist also die Länge $|\vec{C}|$ des Chiralitätsvektors gesucht.

$$d_{\text{SWNT}} = \frac{U_{\text{SWNT}}}{\pi} \implies d_{\text{SWNT}} = \frac{|\vec{C}|}{\pi}$$

Der Gittervektor \vec{a}_1 zeige in Richtung des Einheitsvektors \vec{e}_1 . Dann ergibt sich für \vec{a}_1 und \vec{a}_2 mit a als Gitterkonstante:

$$\vec{a}_1 = \vec{e}_1 \cdot a = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot a \quad \text{und} \quad \vec{a}_2 = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} \cdot a \quad \text{wobei} \quad a = |\vec{a}_1| = |\vec{a}_2|$$

Die Länge des Chiralitätsvektors $|\vec{C}|$ ist über den Typ (n, m) der CNT und die komponentenweise Darstellung gegeben:

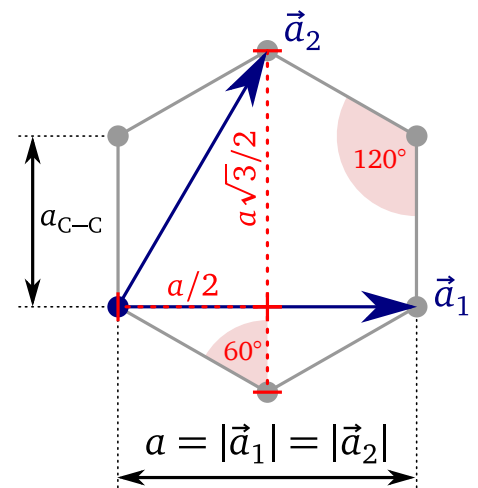
$$\begin{aligned} \vec{C} &= n \cdot \vec{a}_1 + m \cdot \vec{a}_2 \\ \iff \vec{C} &= \begin{pmatrix} n \cdot a_{1,x} + m \cdot a_{2,x} \\ n \cdot a_{1,y} + m \cdot a_{2,y} \end{pmatrix} \\ \iff \vec{C} &= \begin{pmatrix} 1 \cdot n + \frac{1}{2} \cdot m \\ \frac{\sqrt{3}}{2} \cdot m \end{pmatrix} \cdot a \\ |\vec{C}| &= \sqrt{C_x^2 + C_y^2} \\ \iff |\vec{C}| &= a \cdot \sqrt{\left(n + \frac{1}{2}m\right)^2 + \left(\frac{\sqrt{3}}{2}m\right)^2} \\ \iff |\vec{C}| &= a \cdot \sqrt{n^2 + nm + m^2} \end{aligned}$$

Unter Verwendung des Innenwinkels von 120° in einem Sechseck kann die Gitterkonstante a über die Bindungslänge $a_{\text{C-C}}$ im Graphen-Gitter ausgedrückt werden:

$$\sin(60^\circ) = \frac{\frac{a}{2}}{a_{\text{C-C}}} \iff a = \sqrt{3} a_{\text{C-C}}$$

Damit folgt schlussendlich:

$$d_{\text{SWNT}} = \frac{\sqrt{3} a_{\text{C-C}} \cdot \sqrt{n^2 + nm + m^2}}{\pi}$$



Vollständiges Testchip-Layout

Das für diese Arbeit entworfene vollständige Testchip-Layout umfasst insgesamt sechs Ebenen und ist in Abb. 77 dargestellt. In dieser Arbeit wurden nur die ersten drei Ebenen des Layouts und die Finger-Transistoren mit einem globalen *Backgate* genutzt. Neben den tatsächlich verwendeten Strukturen sind verschiedene weitere Transistoren und individuelle *Topgates* für jeden Transistor im Layout vorbereitet.

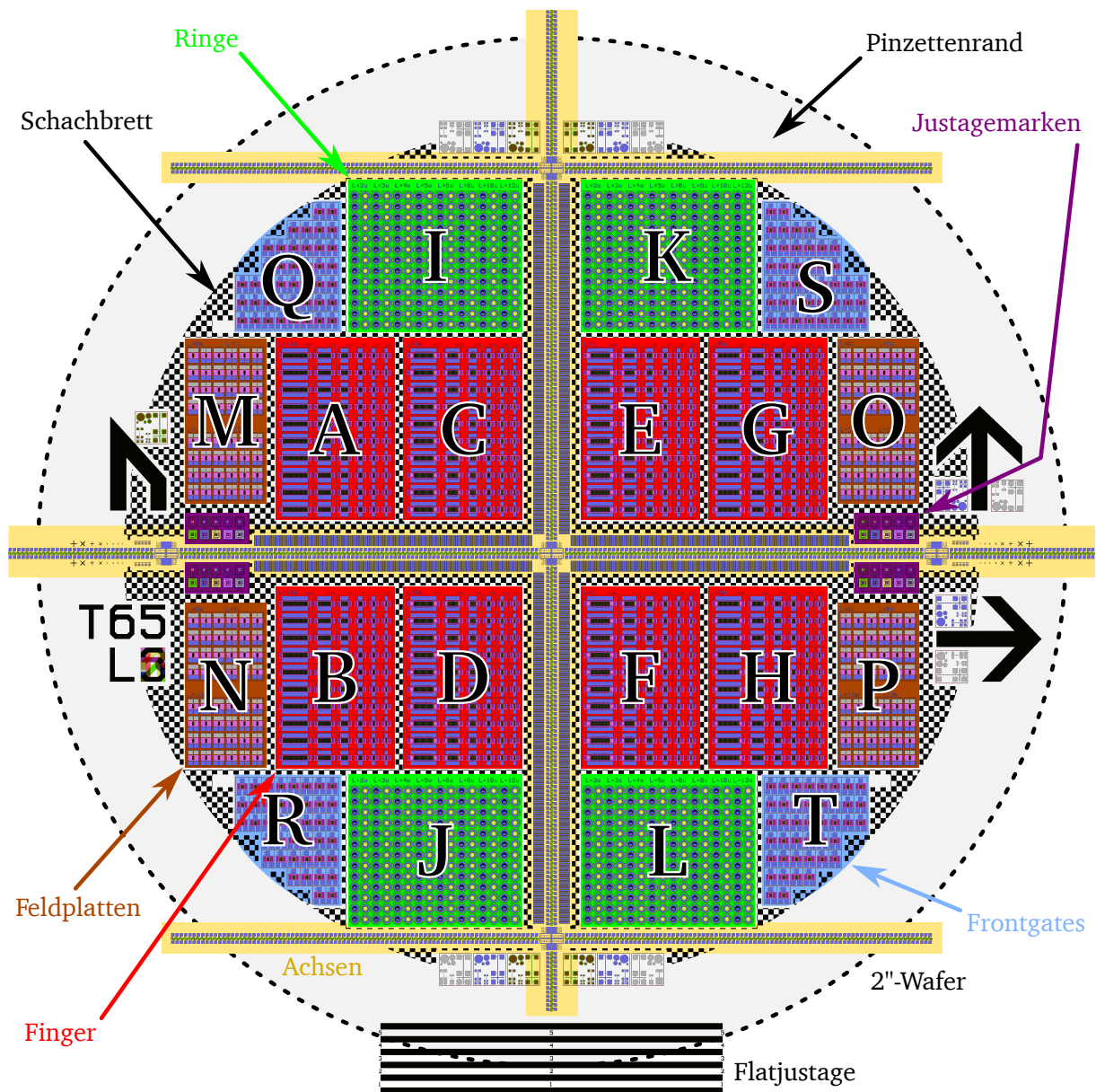


Abb. 77: Vollständiges Testchip-Layout mit sechs Ebenen, weiteren Transistortypen und individuellen *Topgates*.

Die Ebenen des vollständigen Layouts dienen im Einzelnen der Erzeugung der unten genannten Strukturen, wobei die Strukturübertragung für Ebene 1 über nasschemisches Ätzen und für die anderen Ebenen mittels der *Lift off*-Technik erfolgen soll. Neben den in Kapitel 6.1 (Seite 56) näher beschriebenen Finger-Transistoren sind Ring- und Feldplatten-Transistoren und Transistoren mit unterschiedlich breiten *Topgates* in den verschiedenen Feldern des Layouts vorgesehen.

Layout-Ebenen

1. Justagemarken und Gräben zur Abgrenzung der Transistoren gegeneinander
2. strukturierter Katalysator zum Wachsen der CNTs
3. Metallisierung der CNTs mit *Source/Drain*-Kontakten aus Palladium
4. Metallisierung der CNTs mit *Drain*-Kontakten aus einem alternativen Metall
5. *Gate*-Dielektrikum
6. *Topgate*-Kontakt

Finger-Transistoren

- 2x 4 Felder, A bis H, rote Markierungen in Abb. 77
- Felder paarweise identisch: A = B, C = D, E = F und G = H
- vier Kanallängen: 2, 3, 5 und 10 μm
- sechs Kanalweiten: 180, 360, 720, 1440, 2880 und 5760 μm
- *Source/Drain*-Kontakte aus Palladium/Palladium
- *Gate*-Dielektrikum im Kanalbereich
- *Topgates* mit 5 μm *Overlap* auf *Source/Drain*
- 2x 10 Transistoren pro Kombination Kanallänge/Kanalweite, bei 24 verschiedenen Kombinationen
- insgesamt 480 Finger-Transistoren

Ring-Transistoren

- 4 Felder, I bis L, grüne Markierungen in Abb. 77
- alle Felder identisch
- acht Kanallängen: 2, 3, 4, 5, 6, 8, 10 und 12 μm
- eine Kanalweite: 630 μm
- *Source/Drain*-Kontakte aus Palladium/Palladium und aus alternativem Metall/Palladium
- *Gate*-Dielektrikum im Kanalbereich
- *Topgates* mit 5 μm *Overlap* auf *Source/Drain*
- 4x 13 Transistoren pro Kombination Kanallänge/Kontaktmetalle, bei 16 verschiedenen Kombinationen
- insgesamt 832 Ring-Transistoren

Feldplatten-Transistoren

- 4 Felder, M bis P, braune Markierungen in Abb. 77
- alle Felder identisch
- vier Kanallängen: 6, 9, 12 und 15 μm
- eine Kanalweite: 180 μm
- *Source/Drain*-Kontakte aus Palladium/Palladium
- *Gate*-Dielektrikum im Kanalbereich
- für *Source* und *Drain* jeweils eine eigene Feldplatte mit *Overlap* auf den Kanal
- drei *Overlaps*: 1, 2 und 3 μm
- 4x 4 Transistoren pro Kombination Kanallänge/*Overlap*,
bei 12 verschiedenen Kombinationen
- insgesamt 192 Feldplatten-Transistoren

Frontgate-Transistoren

- 4 Felder, Q bis T, blaue Markierungen in Abb. 77
- vier Kanallängen: 15, 20, 25 und 30 μm
- eine Kanalweite: 180 μm
- *Source/Drain*-Kontakte aus Palladium/Palladium, die sich als „Doppeltransistoren“ einen gemeinsamen *Source*-Kontakt teilen
- *Gate*-Dielektrikum im Kanalbereich
- zwei schmale *Topgates* mit einer Breite von 20 und 50 % der Kanallänge
- 42 Transistoren pro Kombination Kanallänge/*Topgate*-Breite,
bei 8 verschiedenen Kombinationen
- insgesamt 336 *Frontgate*-Transistoren

Basisparameter der Herstellung

Sofern nicht anders angegeben und/oder diskutiert wurden folgende Prozessparameter für die jeweiligen Fabrikationsschritte verwendet.

- **Wafer**
 - Hersteller: *Si-Mat*
 - 2'', Cz, (100)-Orientierung, Dicke: 256 bis 306 μm
 - Dotierung: Bor, 0,01–0,02 $\Omega\cdot\text{cm}$
- **Reinigung**
 - „Standard“-Reinigung, 10 min bei 110 °C und 5 min bei Raumtemperatur
 - *piranha*-Reinigung, 15 min
 - *HF-Dip*
 - *RCA clean 1*, 10 min bei 70 °C
 - verdünnter *RCA clean 2*, 2 min bei Raumtemperatur
- **Oxidation**
 - *HF-Dip*
 - Sauerstoff, 9 SE
 - Temperatur: 1000 °C
 - Dauer: 120 min
- **Strukturübertragung**
 - Trocknen der Substrate
 - * über Nacht bei 120 °C in Stickstoff
 - Belacken
 - * Lack: *AZ 5214E* von *MicroChemicals*, 1,4 μm dick
 - * Aufbringen mit Glaskolbenspritze mit Vorsatzfilter
Filter: Teflon, Durchmesser: 25 mm, 0,45 μm -Poren
 - * *spin coater*: 4000 U/min für 30 s
 - * *soft bake*: 110 °C für 50 s
 - Belichten und Entwickeln — für Strukturübertragung durch Ätzen
 - * Kontaktbelichtung
 - * Belichtung: 25 s bei 30 mW/cm² (Weißglaslinse)
 - * Entwicklung: *AZ 726 MIF* von *MicroChemicals*, 1 min

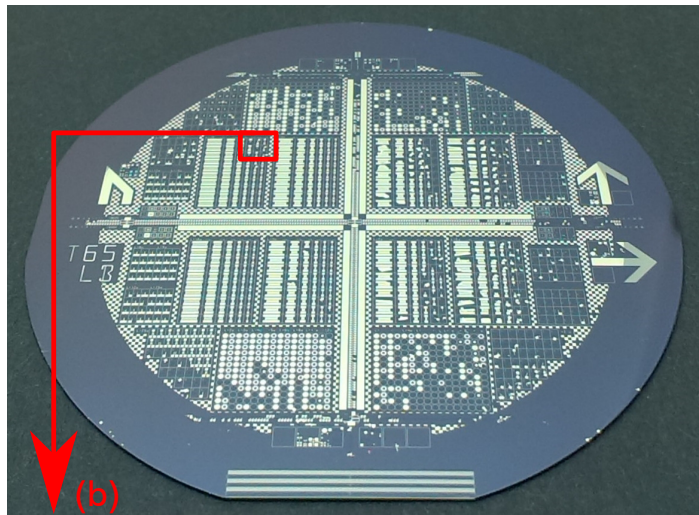
- Belichten und Entwickeln — für Strukturübertragung durch *Lift off*-Technik
 - * Kontaktbelichtung
 - * Belichtung: 5 s bei 15 mW/cm² (Schwarzglaslinse)
 - * *reversal bake*: auf *hot plate* bei 125 °C für 90 s
 - * Flutbelichtung: 2 min
 - * Entwicklung: AZ 726 MIF von MicroChemicals, 3 min
- *Lift off*
 - * Remover: NEP in zwei Bädern bei 70 °C, optional mit Ultraschall
 - * Dauer: jeweils 2 bis 10 min
 - * Spülen in IPA und optional in DI-Wasser
- **Elektronenstrahlverdampfung**
 - Katalysator
 - * 5 nm Aluminium
 - * 1 nm Nickel
 - Kontakte
 - * 10 nm Palladium
 - * 100 nm Nickel
- **Sauerstoff-Plasmabehandlung**
 - Gas: 20 mL/min Sauerstoff, 200 mTorr
 - Leistung: 100 W RF, 0 W DC
 - Dauer: 5 min
- **Tempern & CCVD** (Basisprozess)
 - Spülen
 - * in Inertgas, 16 L/min für 3 min und 8 L/min für 7 min
 - Cluster-Bildung
 - * 5 L/min Inertgas
 - * Temperatur: 870 °C
 - * Dauer: 5 min
 - CNT-Wachstum
 - * 6 L/min Methan und 0,25 L/min Wasserstoff
 - * Temperatur: 870 °C
 - * Dauer: 10 min
 - Kühlen
 - * in Inertgas, 5 L/min für 5 min

Haftungsproblem der Palladium-Kontakte

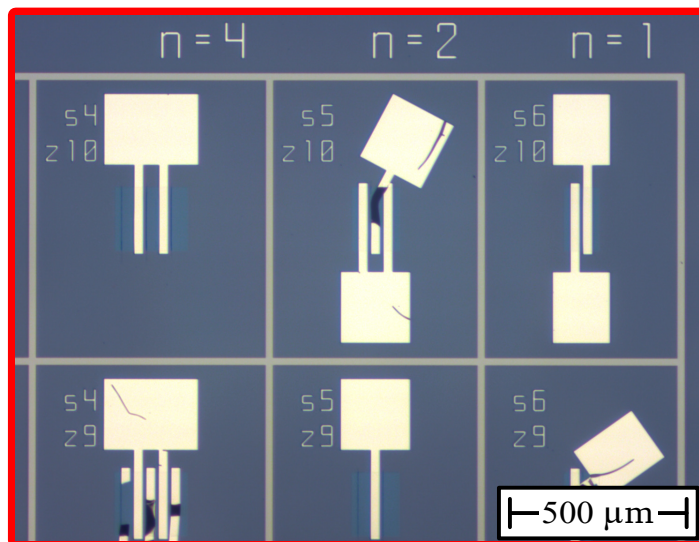
Das in dieser Arbeit als Kontaktmaterial gewählte Palladium hat sich hinsichtlich seiner Prozessierbarkeit mittels *Lift off* als problematisch erwiesen. Zunächst hat sich gezeigt, dass Palladium sehr weich ist, da die Kontakte durch das Aufsetzen der Nadeln bei der elektrischen Charakterisierung stark geschädigt werden. Außerdem sind die Haftungseigenschaften auf Siliziumdioxid unzureichend, da beim *Lift off* nicht nur Bereiche auf Lack entfernt werden, sondern auch solche Bereiche mitgerissen werden, die zurückbleiben sollen. Die Haltbarkeit der Kontakte ließe sich zwar durch die Verwendung sehr dicker Palladium-Schichten verbessern, dies hat aber wiederum das Haftungsproblem weiter verschärft.

Abb. 78a und Abb. 78b zeigen exemplarisch eine Probe, auf der wegen der schlechten Haftung von Palladium ein Großteil der *Source/Drain*-Kontakte durch den *Lift off* zerstört wurden, während Abb. 78c ein gutes Ergebnis zeigt, das mit dem in Kapitel 6.2 (Seite 64) beschriebenen Doppellagen-System unter Verwendung von Nickel als *Capping* erreicht wurde. Die Ablösung der reinen Palladium-Kontakte ist „degenerativ“: Mit jedem neuen Umsetzen der Proben von einem Bad des *Lift off*-Prozesses in das nächste Bad werden mehr und mehr Kontakte zerstört, wobei dickere Palladium-Schichten einen verstärkten Hang zur Ablösung zeigen. Außerdem zeigt sich, dass der abschließende Reinigungsschritt in DI-Wasser, der ursprünglich im Anschluss an eine Spülung in Isopropanol (IPA) verwendet wurde, besonders verheerend für das Endergebnis ist. Die Spülung in DI-Wasser wurde daher in der Folge prinzipiell nicht mehr verwendet.

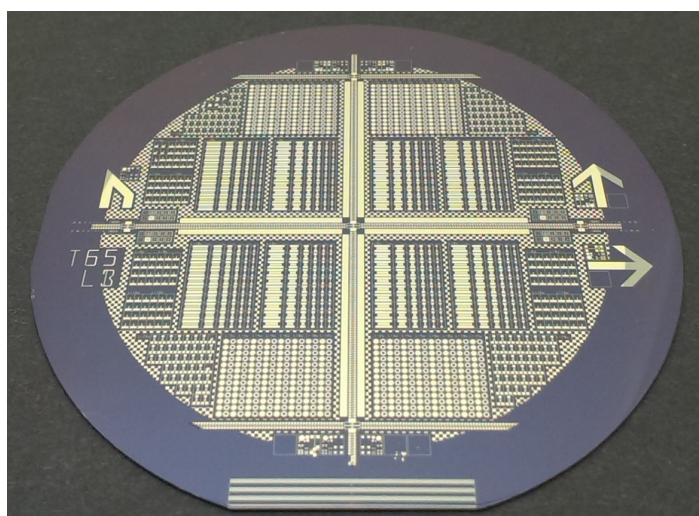
Es ist bekannt, dass die Haftungseigenschaft von Metallen aus der sogenannten Platin-Gruppe auf Oxiden im Allgemeinen und von Palladium auf Siliziumdioxid im Besonderen lediglich bescheidenen sind, da die für eine gute Haftung nötige chemische Wechselwirkung zwischen Schicht und Substrat fehlt [101]. Außerdem ist es wahrscheinlich, dass die Schicht aufgrund der erhöhten Temperatur bei der Bedampfung gegen den Untergrund eine innere Zugspannung aufweist. Schlechte Haftung und innere Spannung haben dann zur Folge, dass sich die Schicht am Rand und besonders an den Ecken der Kontakte leicht anhebt und damit eine Angriffsfläche gegen vorbeiströmende Flüssigkeit bildet. Werden die Wafer dann aus den Becken genommen, werden ausgehend von diesen Angriffsflächen die Kontakte wie Aufkleber von den Wafern abgezogen. Ein solcher Mechanismus erklärt auch, warum das Doppellagen-System die *Lift off*-Ergebnisse verbessert: Durch die geringere Schichtdicke des Palladiums wird zum einen die innere Spannung reduziert, was den Hang zum Anheben der Ecken unterdrückt, zum anderen wird durch die dicke Nickel-*Capping* ein mechanischer Widerstand gegen ein solches Anheben erreicht.



(a) Schlechte Haftung: Kontakte aus reinem Palladium.



(b) Transistoren mit beschädigten Source/Drain-Kontakten.



(c) Gute Haftung: Kontakte mit Nickel-Capping.

Abb. 78: Haftungsproblem der Kontakte.

Wafer-Map zu Kapitel 11.2

Die jeweilige Wafer-Map für die FETs vom Typ A und die FETs vom Typ B aus Kapitel 11.2 ab Seite 108 ist in Abb. 79 und Abb. 80 dargestellt. Die Klassifizierung als defekt/metallisch/halbleitend geschieht gemäß der aus Abb. 69 und Abb. 70 extrahierten und in Tab. 9 zusammengefassten Kennwerte.

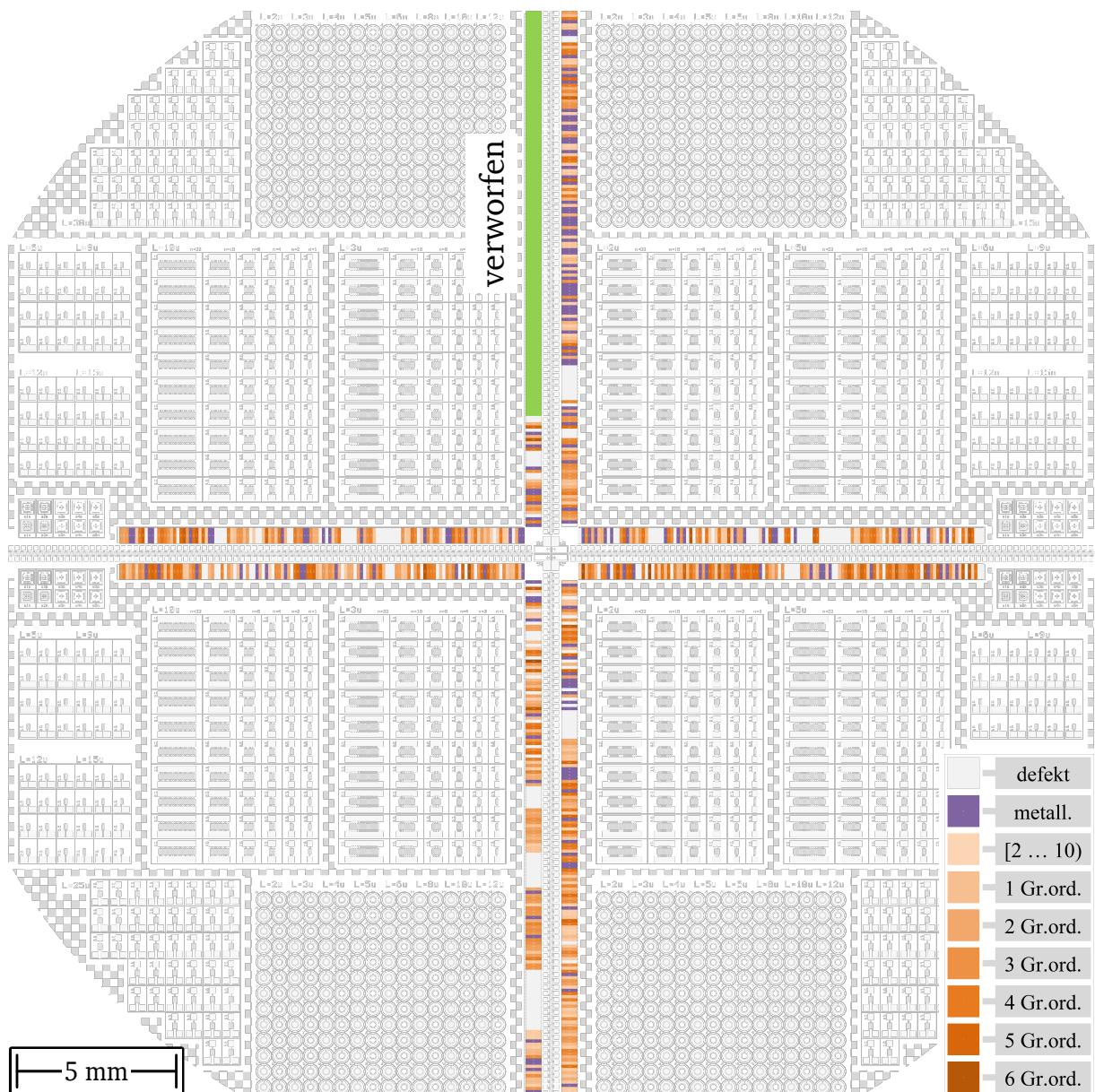
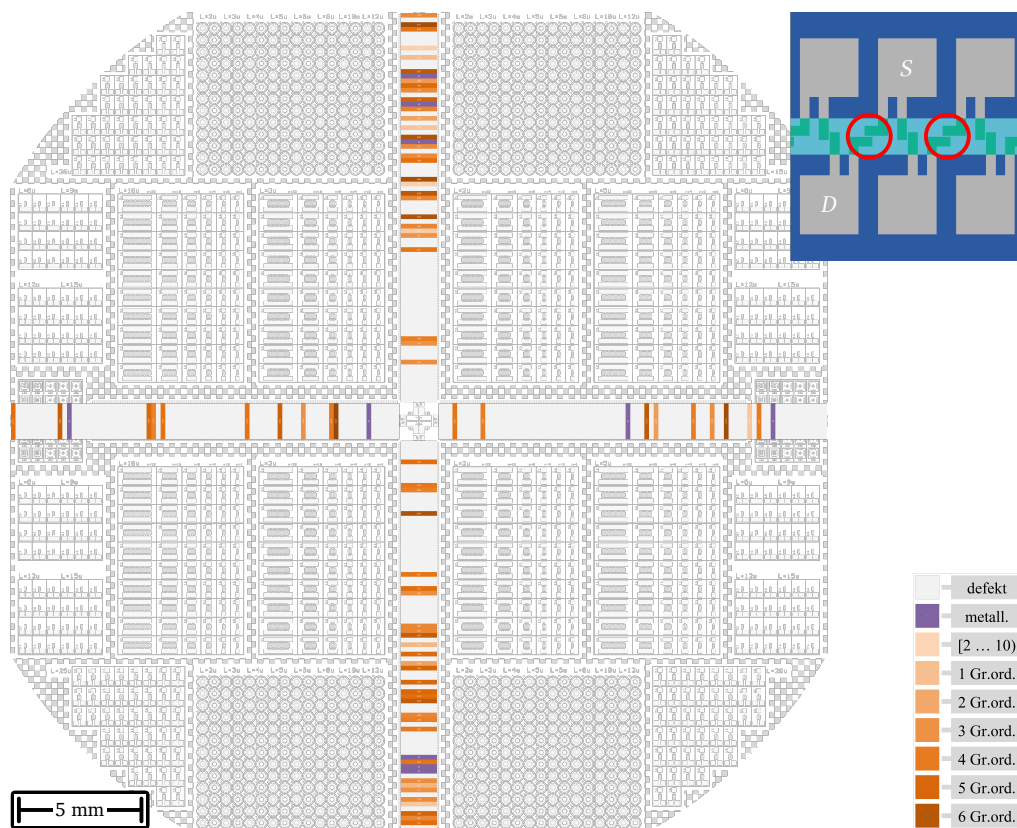
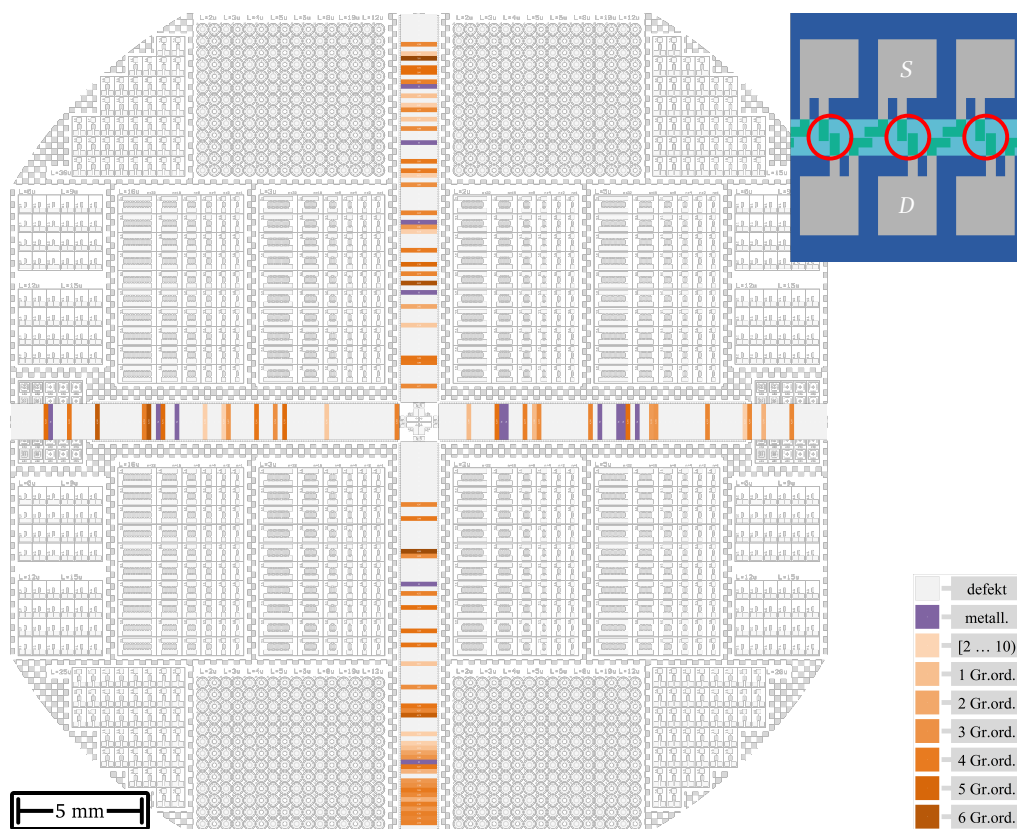


Abb. 79: Wafer-Map für die FETs vom Typ A.



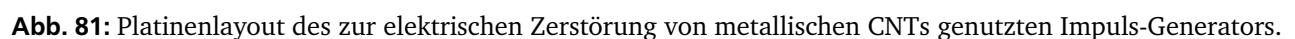
(a) FETs, deren Kanal parallel zum Wafer-Flat liegt



(b) FETs, deren Kanal senkrecht zum Wafer-Flat liegt

Abb. 80: Wafer-Map für die FETs vom Typ B.

Abb. 81 zeigt das Platinenlayout und Abb. 82 den Schaltplan des sogenannten „Impuls-Generators“, der von ANDREAS KRAMER für diese Arbeit entworfen wurde. Die farblichen Markierungen der einzelnen Blöcke und Bezeichnungen der Anschlüsse in Abb. 82 entsprechen der Beschreibung in Kapitel 11.4 (Seite 117) und finden sich auch in Abb. 75 (Seite 118) wieder.



Netzteil

DUT

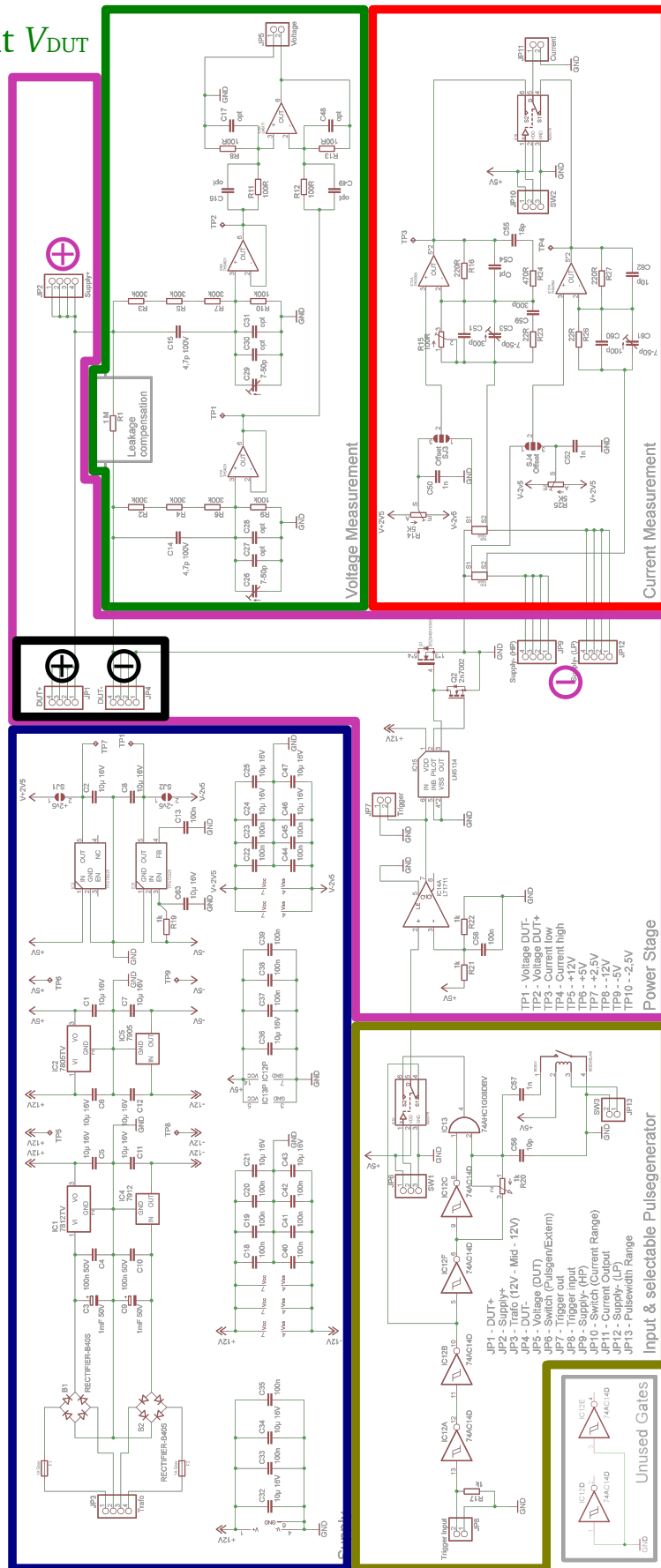
Output V_{DUT}

Output I_{DUT}

Input V_{Supply}

Input t_{Puls}

Abb. 82: Schaltplan des Impuls-Generators.



Verzeichnisse

Abbildungsverzeichnis

1	Entdeckung der CNTs	1
2	Chiralität von CNTs	5
3	Atomorbitale von Kohlenstoff	7
4	Ersten Brillouin-Zone von Graphen	9
5	Konstruktion der ersten Brillouin-Zone von CNTs	11
6	Elektronisches Verhalten von CNTs verschiedener Chiralität	12
7	Verfahren zur Herstellung von CNTs	15
8	Bänderdiagramm eines Metall–Halbleiter-Kontakts (Schottky-Diode)	17
9	Stromfluss durch eine Schottky-Barriere	19
10	Bänderdiagramm eines idealen MOS-Kondensators	21
11	Aufbau und Kennlinien eines Metall–Oxid–Halbleiter-Feldeffekt-Transistor	24
12	Ellipsometer	25
13	Einschichtmodell der Ellipsometrie	26
14	Auswertung ellipsometrischer Messungen	27
15	Rasterkraftmikroskop	29
16	Höhen- und Amplituden-Signal im <i>tapping mode</i> einer AFM-Messung	30
17	Messartefakte in AFM-Messungen	31
18	Messaufbau für die elektrische Charakterisierung	32
19	Reinigungsbäder in der Nasschemie	34
20	Ofenbank	36
21	Grundprinzip der Photolithographie	37
22	Geräte im Gelbraum	38
23	Prinzip der <i>Lift off</i> -Technik	42
24	Anlage zur Elektronenstrahlverdampfung	43
25	CVD-Anlage	45
26	Ablauf einer CVD-Reaktion	46
27	ALD-Zyklus	47
28	ALD-Anlage	48
29	Prinzip der ALD	49
30	Sputteranlage	50

31	RTP-Anlage	51
32	RTP-Störfall: Oberflächlich aufgeschmolzener Wafer	52
33	Trockenätzanlage	53
34	Ozongenerator	53
35	Übersicht des Testchip-Layouts	56
36	Justagemarken des Testchip-Layouts	58
37	Fingertransistoren des Testchip-Layouts	59
38	Prozessablauf zur Herstellung von CNTFETs	60
39	CNTFET-Herstellung, Block 1: <i>Gate-Oxid</i>	61
40	CNTFET-Herstellung, Block 2: Justagemarken	62
41	CNTFET-Herstellung, Block 3: strukturierter Katalysator	63
42	CNTFET-Herstellung, Block 4: Cluster-Bildung	64
43	CNTFET-Herstellung, Block 4: CNT-Wachstum	64
44	CNTFET-Herstellung, Block 5: <i>Source/Drain</i> -Kontakte mit <i>Capping</i>	65
45	CNTFET-Herstellung, Block 6: elektrische Charakterisierung	65
46	Wachstumsmechanismus von CNTs	68
47	Prozessführung in Block 4: Basisprozess nach [6]	70
48	Einfluss der Sauerstoff-Plasmabehandlung des Katalysators (AFM-Messung)	71
49	Prozessführung in Block 4: Dauer des CNT-Wachstums	72
50	Einfluss der Dauer des CNT-Wachstums (AFM-Messung)	73
51	Prozessführung in Block 4: Temperatur der Cluster-Bildung	74
52	Einfluss der Temperatur während der Cluster-Bildung (AFM-Messung)	75
53	Temperatur-Verläufe für verschiedene Inertgase in der CVD-Anlage	77
54	Prozessführung in Block 4: Inertgas während der Cluster-Bildung	78
55	Einfluss des Inertgases während der Cluster-Bildung (AFM-Messung)	79
56	Prozessführung in Block 4: Heizrate der Cluster-Bildung	80
57	Einfluss der Heizrate der Cluster-Bildung (AFM-Messung)	81
58	CNTFETs mit Al/Ni-Katalysator: Prozessführung in Block 4	85
59	Elektrische Charakterisierung von FET 1 und FET 2	88
60	Nominale und effektive Abmessungen des Kanals in einem CNTFET	90
61	Lokalisierung der dielektrischen Durchbrüche (CS-AFM-Messung)	94
62	Querschnitt des Katalysators (TEM-Messung)	95
63	<i>Fencing</i> -Problem bei der Strukturübertragung mittels <i>Lift off</i> -Technik	97
64	Alternative Prozess-Konzepte zur Vermeidung der <i>Gate</i> -Durchbrüche	100
65	Cluster-Bildung und CNT-Wachstum auf den alternativen Schichtstapeln	103

66	CNTFETs mit $\text{Al}_2\text{O}_3/\text{Ni}$ -Katalysator: Prozessführung in Block 4	105
67	CNT-Wachstum für die beiden Katalysatoren im Vergleich (AFM-Messung)	106
68	Flussdiagramm zur Klassifizierung der FETs vom Typ A und der FETs vom Typ B .	109
69	<i>Gate</i> -Ströme und Unterswellenkennlinien der FETs vom Typ A	110
70	<i>Gate</i> -Ströme und Unterswellenkennlinien der FETs vom Typ B	111
71	Klassifizierung der FETs vom Typ A	112
72	Klassifizierung der FETs vom Typ B	113
73	Schematischer Zusammenhang zwischen Ausbeute und nominaler Kanalweite . .	114
74	Elektrische Charakterisierung ausgewählter FETs vom Typ A und FETs vom Typ B .	116
75	Aufbau des Impuls-Generators zur elektrischen Zerstörung metallischer CNTs . . .	118
76	Elektrische Zerstörung von metallische CNTs	119
77	Vollständiges Testchip-Layout	129
78	Haftungsproblem der Kontakte aus reinem Palladium	136
79	Wafer-Map für die FETs vom Typ A	137
80	Wafer-Map für die FETs vom Typ B	138
81	Platinenlayout des Impuls-Generators	139
82	Schaltplan des Impuls-Generators	140

Tabellenverzeichnis

1	Grenzflächenzustände eines idealen MOS-Kondensators	21
2	Verfügbare nominale Kanallängen und -weiten im Testchip-Layout	59
3	Thermische Materialdaten für Stickstoff und Argon	77
4	Cluster-Größen und -Anzahl für verschiedene Heizraten der Cluster-Bildung . . .	82
5	Aus Unterschwellenkennlinien für FET 1, FET 2 und FET R ermittelte Größen	87
6	Vergleich von nominaler und effektiver Kanallänge bzw. -weite	91
7	Alternative Schichtstapel zur Vermeidung der <i>Gate</i> -Durchbrüche	101
8	Nominale Kanallängen und -weiten und Kontakt-Flächen der FETs vom Typ A und der FETs vom Typ B	109
9	Grenzwerte der statistischen Untersuchung an den FETs vom Typ A und den FETs vom Typ B	109

Abkürzungsverzeichnis

AFM	Rasterkraftmikroskopie <i><u>a</u>tom<u>i</u>c <u>f</u>orce <u>m</u>icrosco<u>p</u>y</i>
ALD	Atomlagenabscheidung <i><u>a</u>tom<u>i</u>c <u>l</u>ayer <u>d</u>eposition</i>
B	<i>Bulk</i> – Substrat-Anschluss
BOE	gepufferte Oxid-Ätze <i><u>b</u>uffered <u>o</u>xide <u>e</u>tch</i>
CCVD	katalytisch-chemische Gasphasenabscheidung <i><u>c</u>atalytic <u>c</u>hemical <u>v</u>apour <u>d</u>eposition</i>
CDE	chemisches Trockenätzen <i><u>c</u>hemical <u>d</u>ry <u>e</u>tching</i>
CNT	Kohlenstoffnanoröhre <i><u>c</u>arbon <u>n</u>anot<u>u</u>be</i>
CNTFET	Kohlenstoffnanoröhren-Feldeffekt-Transistor <i><u>c</u>arbon <u>n</u>anot<u>u</u>be <u>f</u>ield-<u>e</u>ffect <u>t</u>ransistor</i>
CS-AFM	Strom-Spannungs-Mikroskopie <i><u>c</u>urrent <u>s</u>ensing <u>a</u>tom<u>i</u>c <u>f</u>orce <u>m</u>icrosco<u>p</u>y</i>
CVD	chemische Gasphasenabscheidung <i><u>c</u>hemical <u>v</u>apour <u>d</u>eposition</i>
D	<i>Drain</i> – (Signal-)Kontakt eines FET
DI-Wasser	deionisiertes Wasser
DUT	Prüfling / zu prüfendes Bauelement <i><u>d</u>evice <u>u</u>nder <u>t</u>est</i>
FET	Feldeffekt-Transistor <i><u>f</u>ield-<u>e</u>ffect <u>t</u>ransistor</i>
G	<i>Gate</i> – (Steuer-)Kontakt eines FET

IHTN	INSTITUT FÜR HALBLEITERTECHNIK UND NANOELEKTRONIK
IPA	Isopropanol
MOS	Metall–Oxid–Halbleiter <i>metal–oxide–semi-conductor</i>
MOSFET	Metall–Oxid–Halbleiter-Feldeffekt-Transistor <i>metal–oxide–semi-conductor field-effect transistor</i>
MWNT	mehrwandige Kohlenstoffnanoröhre <i>multi-walled nanotube</i>
NEP	n-Ethyl-2-pyrrolidon
PVD	physikalische Gasphasenabscheidung <i>physical vapour deposition</i>
RTP	Kurzzeit-Prozessierung <i>rapid thermal processing</i>
S	Source – (Signal-)Kontakt eines FET
SB	Schottky-Barriere
SEM	Rasterelektronenmikroskopie <i>scanning electron microscopy</i>
SWNT	einwandige Kohlenstoffnanoröhre <i>single-walled nanotube</i>
TEM	Transmissionselektronenmikroskopie <i>transmission electron microscopy</i>
TMA	Trimethylaluminium

Symbolverzeichnis

a	Gitterkonstante im Graphen-Gitter
\vec{a}_1	Gittervektor im Graphen-Gitter
\vec{a}_2	Gittervektor im Graphen-Gitter
a_{C-C}	Kohlenstoff-Kohlenstoff-Bindungslänge in Graphen
A_{Pad}	nominale Fläche eines <i>Source/Drain</i> -Kontakts eines FET im Testchip-Layout
\vec{C}	Chiralitätsvektor einer CNT
d_{Oxid}	Oxiddicke
d_{SWNT}	Durchmesser einer SWNT
E_0	Vakuumenergieniveau
E_C	Leitungsbandkante
E_f	Fermi-Energie bzw. Ferminiveau
$E_{f,M}$	Fermi-Energie eines Metalls
$E_{f,S}$	Fermi-Energie eines Halbleiters
E_g	Bandlücke eines Halbleiters oder einer CNT
E_i	intrinsisches Energieniveau
E_{Puls}	Energie des Überlastungs-Impulses
E_V	Valenzbandkante
I_{DUT}	bei der Impuls-Überlastung fließender Strom I_{DS}
I_{DS}	Strom, der zwischen <i>Source</i> und <i>Drain</i> fließt
$I_{DS, \text{sat}}$	Sättigungsstrom eines FET
I_G	Strom, der durch das <i>Gate</i> fließt
<i>Inert</i>	Inertgas, das während der Cluster-Bildung genutzt wird
I_{Oxid}	Leckstrom, der durch das <i>Gate-Oxid</i> fließt

I_{off}	Strom, der zwischen <i>Source</i> und <i>Drain</i> fließt im <i>Off</i> -Zustand
I_{on}	Strom, der zwischen <i>Source</i> und <i>Drain</i> fließt im <i>On</i> -Zustand
I_{SB}	Strom, der durch eine Schottky-Barriere fließt
\vec{k}	Wellenvektor im reziproken Raum
k_{B}	Boltzmann-Konstante
$L_{\text{Kanal nom}}$	nominale Kanallänge eines FET im Testchip-Layout
n	Anzahl der Fingerpaare in einem Fingertransistor (sofern nicht anders definiert)
(n, m)	Chiralitätstyp einer CNT
<i>Off</i>	sperrender Zustand eines FET
<i>On</i>	leitender Zustand eines FET
<i>On/Off</i>	Verhältnis von I_{on} zu I_{off}
P_{Puls}	nominale <i>Peak</i> -Leistung des Überlastungs-Impulses
R_{DUT}	initialer Widerstand des Bauelements bei der Impuls-Überlastung
R_{off}	elektrischer Widerstand im <i>Off</i> -Zustand
R_{on}	elektrischer Widerstand im <i>On</i> -Zustand
R_{cool}	Kühlrate nach der Cluster-Bildung oder dem CNT-Wachstum
R_{Temp}	Heizrate der Cluster-Bildung
S_{subthr}	Steigung der Unterschwellenkennlinie am Übergang vom <i>Off</i> - zum <i>On</i> -Zustand
\vec{T}	Translationsvektor einer CNT
t_{CCVD}	Dauer des CNT-Wachstums
T_{CCVD}	Temperatur des CNT-Wachstums
t_{Puls}	nominale Dauer des Überlastungs-Impulses
t_{Temp}	Dauer der Cluster-Bildung
T_{Temp}	Temperatur der Cluster-Bildung
U_{SWNT}	Umfang einer CNT
V_{bi}	eingebautes Potential
V_{DS}	Spannung, die zwischen <i>Source</i> und <i>Drain</i> anliegt

$V_{DS, sat}$	Sättigungsspannung eines MOSFET
V_{GS}	Spannung, die am <i>Gate</i> anliegt
V_{DUT}	bei der Impuls-Überlastung anliegende Spannung V_{DS}
V_{Supply}	Versorgungsspannung des Impuls-Generators
V_T	Schwellspannung bzw. Einsatzspannung
$W_{\text{Kanal nom}}$	nominale Kanalweite eines FET im Testchip-Layout
W_{RLZ}	Weite der Raumladungszone
φ_{SB}	Höhe der SB
φ_M	Austrittsarbeit eines Metalls
φ_S	Austrittsarbeit eines Halbleiters
χ_S	Elektronenaffinität eines Halbleiters

Farbcodierung

Sofern möglich wird in den schematischen Darstellungen dieser Arbeit für die verschiedenen Materialien und Elemente die folgende Farbgebung eingehalten:

Aluminium

Aluminiumoxid

Kohlenstoff

Kohlenstoffnanoröhre

Palladium

photosensitiver Lack

lithographische Maske

Nickel

thermisches Oxid

Siliziumnitrid

Sauerstoff

Schicht (allgemein)

(Silizium-)Substrat

Wasserstoff

Literaturverzeichnis

- [1] J. S. Kilby, *Invention of the integrated circuit*, IEEE Transactions on Electron Devices 23 (1976) 648–654
- [2] J. S. Kilby, *Miniaturized electronic circuits*, US Patent 3.138.743 (1964)
- [3] G. E. Moore, *Cramming more components onto integrated circuits*, Electronics 38 (1965) 114–116
- [4] F. Wessely, *CMOS ohne Dotierstoffe: Neuartige siliziumbasierte Nanodraht-Feldeffekt-Bauelemente*, Dissertation, 2011, Institut für Halbleitertechnik und Nanoelektronik, Technische Universität Darmstadt
- [5] A. Heinzig et al., *Reconfigurable silicon nanowire transistors*, Nano Letters 12 (2012) 119–124
- [6] L. Rispal, *Large scale fabrication of field-effect devices based on in situ grown carbon nanotubes*, Dissertation, 2009, Institut für Halbleitertechnik und Nanoelektronik, Technische Universität Darmstadt
- [7] P. J. Wessely, *Graphen-Transistoren: Silizium CMOS kompatible Herstellung für Anwendungen in der Nanoelektronik*, Dissertation, 2013, Institut für Halbleitertechnik und Nanoelektronik, Technische Universität Darmstadt
- [8] L. V. Radushkevich und V. M. Lukyanovich, О структуре углерода, образующегося при термическом разложении окиси углерода на железном контакте [*About the structure of carbon formed by thermal decomposition of carbon monoxide on iron substrate*], Soviet Journal of Physical Chemistry 26 (1952) 88–95
- [9] S. Iijima, *Helical microtubules of graphitic carbon*, Nature 354 (1991) 56–58
- [10] S. J. Tans, A. R. M. Verschueren und C. Dekker, *Room-temperature transistor based on a single carbon nanotube*, Letters to Nature 393 (1998) 49–52
- [11] O. Vohler et al., *Carbon, 2. Diamond*, in: Ullmann's Encyclopedia of Industrial Chemistry, Wiley, 2010
- [12] K. S. Novoselov et al., *Electric field effect in atomically thin carbon films*, Science 306 (2004) 666–669

-
- [13] V.I. Kasatochkin et al., *Crystalline forms of linear modification of carbon*, Doklady Akademii Nauk SSSR 177 (1967) 358–360
- [14] M. Liu et al., *Carbyne from first principles: Chain of C atoms, a nanorod or a nanorope*, ACS Nano 7 (2013) 10075–10082
- [15] P.P.K. Smith und P.R. Buseck, *Carbyne forms of carbon: Do they exist?*, Science 216 (1982) 984–986
- [16] B. Pan et al., *Carbyne with finite length: The one-dimensional*, Science Advances 1 (2015) 1–10
- [17] H.W. Kroto et al., *C₆₀: Buckminsterfullerene*, Nature 318 (1985) 162–163
- [18] S. Sawada und N. Hamada, *Energetics of carbon nano-tubes*, Solid State Communications 83 (1992) 917–919
- [19] E. T. Thostenson, Z. Ren und T.-W. Chou, *Advances in the science and technology of carbon nanotubes and their composites: A review*, Composites Science and Technology 61 (2001) 1899–1921
- [20] N. Hamada, S.-I. Sawada und A. Oshiyama, *New one-dimensional conductors: Graphitic microtubules*, Physical Review Letters 68 (1992) 1579–1581
- [21] S. Reich, Ch. Thomsen und J. Maultzsch, *Carbon nanotubes: Basic concepts and physical properties*, Wiley, 2004
- [22] A. H. Castro Neto et al., *The electronic properties of graphene*, Reviews of Modern Physics 81 (2009) 109–162
- [23] N. W. Ashcroft und N. D. Mermin, *Solid state physics*, Harcourt College Publishers, 1976
- [24] Chemiezauber.de, *Kohlenstoff und Kohlenstoffverbindungen*, www.chemiezauber.de, abgerufen am 23.01.2018
- [25] P.R. Wallace, *The band theory of graphite*, Physical Review 71 (1947) 622–634
- [26] R. Saito, G. Dresselhaus und M. S. Dresselhaus, *Trigonal warping effect of carbon nanotubes*, Physical Review B 61 (2000) 2981–2990
- [27] Maruyama-Chiashi Laboratory, *Kataura-plot for resonant raman: Discussions of 1D DOS*, www.photon.t.u-tokyo.ac.jp, abgerufen am 23.01.2018
- [28] S. Blatt, *Dielectrophoresis of single-walled carbon nanotubes*, Dissertation, 2008, Institut für Nanotechnologie, Universität Karlsruhe
- [29] J.-C. Charlier, X. Blase und S. Roche, *Electronic and transport properties of nanotubes*, Reviews of Modern Physics 79 (2007) 677–732

-
- [30] A. Gohier et al., *Carbon nanotube growth mechanism switches from tip- to base-growth with decreasing catalyst particle size*, *Carbon* 46 (2008) 1331–1338
- [31] R. Seidel et al., *Chemical vapor deposition growth of single-walled carbon nanotubes at 600 °C and a simple growth model*, *Journal of Physical Chemistry B* 108 (2004) 1888–1893
- [32] R. Seidel et al., *In-Situ contacted single-walled carbon nanotubes and contact improvement by electroless deposition*, *Nano Letters* 3 (2003) 965–968
- [33] R. Seidel, *Carbon nanotube devices*, Dissertation, 2004, Fakultät Maschinenwesen, Technischen Universität Dresden
- [34] D. S. Bethune et al., *Cobalt-catalysed growth of carbon nanotubes with single-atomic-layer walls*, *Nature* 363 (1993) 605–607
- [35] C. Journet und P. Bernier, *Production of carbon nanotubes*, *Applied Physics A: Materials Science & Processing* 67 (1998)
- [36] V. Popov, *Carbon nanotubes: Properties and application*, *Materials Science and Engineering: R: Reports* 43 (2004) 61–102
- [37] M. S. Dresselhaus, G. Dresselhaus und Ph. Avouris, *Carbon nanotubes — Synthesis, structure, properties, and applications*, Springer, 2001
- [38] R. F. Pierret, *Semiconductor device fundamentals*, Addison-Wesley, 1996
- [39] S. M. Sze und K. K. Ng, *Physics of semiconductor devices*, Wiley-Interscience, 3rd Edition, 2007
- [40] D. K. Schroder, *Semiconductor material and device characterization*, Wiley-Interscience, 3. Auflage, 2006
- [41] H. Fujiwara, *Spectroscopic ellipsometry — Principles and applications*, Wiley, Japanese Edition, 2007
- [42] Gwyddion — *Free SPM data analysis software*, www.gwyddion.net, Version 2.47, abgerufen am 18.11.2016
- [43] S. N. Magonov und M.-H. Whangbo, *Scanning probe microscopes*, in: *Surface analysis with STM and AFM: Experimental and theoretical aspects of image analysis*, Wiley-VCH, 1996
- [44] NanoAndMore, *AFM-Spitze „Tap300Al-G“ von BudgetSensors*, www.nanoandmore.com, abgerufen am 23.01.2018
- [45] G. Binnig, C. F. Quate und C. Gerber, *Atomic force microscope*, *Physical Review Letters* 56 (1986) 930–933

-
- [46] G. Haugstad, *Atomic force microscopy: Understanding basic modes and advanced applications*, Wiley, 2012
- [47] K. Wapner, *Grenzflächenchemische und elektrochemische Untersuchungen zur Haftung und Enthftung an modifizierten Klebstoff/Metall-Grenzflächen*, Dissertation, 2006, Fakultät für Chemie, Ruhr-Universität Bochum
- [48] Y. Stefanov, *The application of atomic force microscopy in semiconductor technology — Towards high-k gate dielectric integration*, Dissertation, 2012, Institut für Halbleitertechnik und Nanoelektronik, Technische Universität Darmstadt
- [49] W. Kern, *Handbook of semiconductor wafer cleaning technology: Science, technology and applications — Materials science and process technology series*, Noyes Publications, 1993
- [50] C. Cowache, P. Boelen und I. Kashkoush, *Evaluation of advanced pre-gate cleanings*, Sixth International Symposium on Cleaning Technology in Semiconductor Device Manufacturing, 17.–22.10.1999, Honolulu, HI, USA
- [51] W. Kern, *The evolution of silicon wafer cleaning technology*, Journal of The Electrochemical Society 137 (1990) 1887–1892
- [52] R. Webb, R. Glahn und R. S. Ridley, Sr., *Effects of dilute HCl wafer cleaning solutions on borophosphosilicate glass films*, Advanced Semiconductor Manufacturing Conference and Workshop, 12.–14.09.2000, Boston, MA, USA
- [53] R. C. Jaeger, *Introduction to microelectronic fabrication*, Prentice Hall, 2nd Edition, 2001
- [54] B. E. Deal und A. S. Grove, *General relationship for the thermal oxidation of silicon*, Journal of Applied Physics 36 (1965) 3770–3778
- [55] W. Langheinrich, *Technologie der integrierten Schaltungen*, Scriptum zur Vorlesung, 2. Auflage, 1997, Institut für Halbleitertechnik, Technische Universität Darmstadt
- [56] C. Mack, *Fundamental principles of optical lithography — The science of microfabrication*, Wiley, 2007
- [57] S. Wolf und R. N. Tauber, *Silicon processing for the VLSI era, Volume 1 — Process technology*, Lattice Press, 1986
- [58] MicroChemicals, *Lift off-Prozesse mit Fotolacken*, Anwendungsinformation, 2013
- [59] D. Widmann, H. Mader und H. Friedrich, *Technology of integrated circuits*, Springer, 2000
- [60] G. Sauerbrey, *Verwendung von Schwingquarzen zur Wägung dünner Schichten und zur Mikrowägung*, Zeitschrift für Physik 155 (1959) 206–222
- [61] K. L. Choy, *Chemical vapour deposition of coatings*, Progress in Materials Science 48 (2003) 57–170

-
- [62] H. O. Pierson, *Handbook of chemical vapor deposition: Principles, technology and applications*, Noyes Publications, 2000
- [63] T. Suntola und J. Antson, *Atomic layer epitaxy*, US Patent 4.058.430 (1977)
- [64] T. Mikolajick, *Technologien der Mikro- und Nanoelektronik*, Vorlesung, Wintersemester 2008/09, 2008, Institut für Elektronik- und Sensormaterialien, Technische Universität Bergakademie Freiberg
- [65] A. C. Dillon et al., *Surface chemistry of Al_2O_3 deposition using $\text{Al}(\text{CH}_3)_3$ and H_2O in a binary reaction sequence*, *Surface Science* 322 (1995) 230–242
- [66] R. L. Puurunen, *Surface chemistry of atomic layer deposition: A case study for the trimethylaluminum/water process*, *Journal of Applied Physics* 97 (2005) 121301
- [67] J.-F. Fan und K. Toyoda, *Self-limiting behavior of the growth of Al_2O_3 using sequential vapor pulses of TMA and H_2O_2* , *Applied Surface Science* 60/61 (1992) 765–769
- [68] H. Kumagai und K. Toyoda, *In situ ellipsometric diagnostics for controlled growth of metal oxides with surface chemical reactions*, *Applied Surface Science* 82–83 (1994) 481–486
- [69] L. Hiltunen et al., *Growth and characterization of aluminium oxide thin films deposited from various source materials by atomic layer epitaxy and chemical vapor deposition processes*, *Materials Chemistry and Physics* 28 (1991) 379–388
- [70] A. Bogaerts et al., *Gas discharge plasmas and their applications*, *Spectrochimica Acta Part B: Atomic Spectroscopy* 57 (2002) 609–658
- [71] M. L. Green et al., *Ultrathin ($< 4\text{ nm}$) SiO_2 and Si-O-N gate dielectric layers for silicon microelectronics: Understanding the processing, structure, and physical and electrical limits*, *Journal of Applied Physics* 90 (2001) 2057–2121
- [72] P. J. Timans, *Rapid thermal processing technology for the 21st century*, *Materials Science in Semiconductor Processing* 1 (1998) 169–179
- [73] AST Elektronik GmbH, *System operation SHS hardware manual*, Handbuch, 1995
- [74] A. Javey et al., *Ballistic carbon nanotube field-effect transistors*, *Letters to Nature* 424 (2003) 654–657
- [75] D. Widmann, *Metallization for integrated circuits using a lift-off technique*, *IEEE Journal of Solid-State Circuits* 11 (1976) 466–417
- [76] J. P. Tessonier und D. S. Su, *Recent progress on the growth mechanism of carbon nanotubes: A review*, *ChemSusChem* 4 (2011) 824–47
- [77] S. Hofmann et al., *Surface diffusion: The low activation energy path for nanotube growth*, *Physical Review Letters* 95 (2005) 036101

-
- [78] S. A. Moshkalyov et al., *Carbon nanotubes growth by chemical vapor deposition using thin film nickel catalyst*, *Materials Science and Engineering: B* 112 (2004) 147–153
- [79] PeriodenSystem.info, Einträge „Stickstoff“ und „Argon“, www.periodensystem.info, abgerufen am 03.01.2017
- [80] COMSOL, *Multiphysics*, User's Guide, 2012
- [81] W. Ostwald, *Über die vermeintliche Isomerie des roten und gelben Quecksilberoxyds und die Oberflächenspannung fester Körper*, *Zeitschrift für Physikalische Chemie* 34 (1900) 495–503
- [82] M. Hillert, *On the theory of normal and abnormal grain growth*, *Acta Metallurgica* 13 (1965) 227–238
- [83] R. A. Oriani, *Ostwald ripening of precipitates in solid matrices*, *Acta Metallurgica* 12 (1965) 1399–1409
- [84] J. Kong et al., *Quantum interference and ballistic transmission in nanotube electron waveguides*, *Physical Review Letters* 87 (2001) 106801
- [85] W. Hoenlein et al., *Carbon nanotubes for microelectronics: Status and future prospects*, *Materials Science and Engineering C* 23 (2003) 663–670
- [86] M. S. Fuhrer et al., *Crossed nanotube junctions*, *Science* 288 (2000) 494–496
- [87] S. Kremmer et al., *Characterization of silicon gate oxides by conducting atomic force microscopy*, *Surface and Interface Analysis* 33 (2002) 168–172
- [88] M. Porti et al., *Systematic characterization of soft- and hard-breakdown spots using techniques with nanometer resolution*, *Microelectronic Engineering* 84 (2007) 1956–1959
- [89] Siemens AG, Corporate Technology, *Untersuchung planarer Kohlenstofffilme auf Silizium*, Messung als Auftragsarbeit, 2010
- [90] E. Philofsky und E. Hall, *A Review of the limitations of aluminum thin films on semiconductor devices*, *IEEE Transactions on Parts, Hybrids and Packaging* 11 (1975) 281–290
- [91] J. R. Black, *Metallization failures in integrated circuits*, Technical Report, 1968, Semiconductor Products Division of Motorola
- [92] G. Schnable und R. S. Keen, *Aluminum metallization — Advantages and limitations for integrated circuit applications*, *Proceedings of the IEEE* 57 (1969) 1570–1580
- [93] L. Rispal et al., *Self-aligned fabrication process based on sacrificial catalyst for Pd-contacted carbon nanotube field-effect transistors*, *ECS Transactions* 11 (2007) 53–61

-
- [94] B. Subramanian, V. Swaminathan und M. Jayachandran, *Micro-structural and optical properties of reactive magnetron sputtered Aluminum Nitride (AlN) nanostructured films*, *Current Applied Physics* 11 (2011) 43–49
- [95] K. Koski, J. Hölsä und P. Juliet, *Voltage controlled reactive sputtering process for aluminium oxide thin films*, *Thin Solid Films* 326 (1998) 189–193
- [96] S. Ramanathan et al., *Growth and characterization of ultrathin zirconia dielectrics grown by ultraviolet ozone oxidation*, *Applied Physics Letters* 79 (2001) 2621–2623
- [97] E. Pop et al., *Electrical and thermal transport in metallic single-wall carbon nanotubes on insulating substrates*, *Journal of Applied Physics* 101 (2007) 093710
- [98] M. Shrivastava, N. Kulshrestha und H. Gossner, *ESD investigations of multiwalled carbon nanotubes*, *IEEE Transactions on Device and Materials Reliability* 14 (2014) 555–563
- [99] P.G. Collins et al., *Current saturation and electrical breakdown in multiwalled carbon nanotubes*, *Physical Review Letters* 86 (2001) 3128–3131
- [100] S. Wang et al., *High-field electrical transport and breakdown behavior of double-walled carbon nanotube field-effect transistors*, *Carbon* 45 (2007) 760–765
- [101] K. G. Kreider und F. DiMeo, *Platinum/palladium thin-film thermocouples for temperature measurements on silicon wafers*, *Sensors and Actuators A* 69 (1998) 46–52

Eigene Veröffentlichungen

Publikationen

Udo Schwalke, Pia Juliane Wessely, Frank Wessely, Martin Keyn und Lorraine Rispal, *Nanoelectronics: From silicon to graphene*, [IEEE Proceedings of the 7th International Conference on Design and Technology of Integrated Systems in Nanoscale Era \(2012\)](#) 1–3

Martin Keyn, Frank Wessely und Udo Schwalke, *Feasibility study on in situ CCVD grown CNTs for field-effect power device applications*, [IEEE Proceedings of the 7th International Conference on Design and Technology of Integrated Systems in Nanoscale Era \(2012\)](#) 1–3

Martin Keyn und Udo Schwalke, *Multi-CNTFETs for power device applications: Investigation of CCVD grown CNTs by means of atomic force microscopy*, [IEEE Proceedings of the 8th International Conference on Design and Technology of Integrated Systems in Nanoscale Era \(2013\)](#) 1–5

Martin Keyn, Andreas Kramer und Udo Schwalke, *Dependence of annealing temperature on cluster formation during in situ growth of CNTs*, [IEEE Proceedings of the 9th International Conference on Design and Technology of Integrated Systems in Nanoscale Era \(2014\)](#) 1–5

Martin Keyn und Udo Schwalke, *Formation process of nickel nano-clusters for catalytic growth of carbon nanotubes for use in field-effect transistors*, [ECS Transactions 64 \(2015\)](#) 1–11

Martin Keyn, Tillmann Krauss, Andreas Kramer und Udo Schwalke, *Evaluation of different high κ materials for in situ growth of carbon nanotubes*, [ECS Transactions 75 \(2016\)](#) 65–71

Konferenzbeiträge

Vorträge

Udo Schwalke, Pia Juliane Wessely, Frank Wessely, Emrah Birinci, Martin Keyn und Lorraine Rispal, *In situ CCVD growth of hexagonal carbon for CMOS-compatible nanoelectronics: From nanotube field-effect devices to graphene transistors*, 12th Trends in Nanotechnology International Conference, 21.–25.11.2011, Teneriffa, Kanarische Inseln, Spanien

Udo Schwalke, Pia Juliane Wessely, Frank Wessely, Martin Keyn und Lorraine Rispal, *Nanoelectronics: From silicon to graphene*, 7th International Conference on Design and Technology of Integrated Systems in Nanoscale Era, 16.–18.05.2012, Gammarth, Tunesien

Martin Keyn und Udo Schwalke, *Multi-CNTFETs for power device applications: Investigation of CCVD grown CNTs by means of atomic force microscopy*, 8th International Conference on Design and Technology of Integrated Systems in Nanoscale Era, 26.–28.03.2013, Abu Dhabi, Vereinigte Arabische Emirate

Martin Keyn, Andreas Kramer und Udo Schwalke, *Dependence of annealing temperature on cluster formation during in situ growth of CNTs*, 9th International Conference on Design and Technology of Integrated Systems in Nanoscale Era, 06.–08.05.2014, Santorin, Griechenland

Martin Keyn und Udo Schwalke, *Highly parallelized carbon nanotubes for field-effect device applications*, 226th Meeting of the Electrochemical Society, 05.–10.10.2014, Cancún, Mexiko

Martin Keyn, Tillmann Krauss, Andreas Kramer und Udo Schwalke, *Evaluation of different high κ materials for in situ growth of carbon nanotubes*, PRiME 2016 / 230th Meeting of the Electrochemical Society, 02.–07.10.2016, Honolulu, Hawaii, USA

Poster

Martin Keyn, Pia Juliane Wessely, Frank Wessely, Lorraine Rispal, Johannes Palm und Udo Schwalke, *Feasibility study on in situ CCVD grown CNTs for field-effect power device applications*, 7th International Conference on Design and Technology of Integrated Systems in Nanoscale Era, 16.–18.05.2012, Gammarth, Tunesien

Martin Keyn, Frank Wessely und Udo Schwalke, *Large-scale parallelization of in situ CCVD grown carbon nanotubes for power devices*, Junior Euromat 2012, 23.–27.07.2012, Lausanne, Schweiz

Andere Arbeiten

Martin Keyn, *Atomic Layer Deposition in der Mikro- und Nanoelektronik*, Literaturarbeit, 2008, Institut für Elektronik- und Sensormaterialien, Technische Universität Bergakademie Freiberg

Martin Keyn, *Untersuchungen zur Vermeidung von Zwillingsbildungen bei der Züchtung von Galliumarsenid-Kristallen*, Ingenieursarbeit, 2009, Freiburger Compound Materials GmbH und Institut für Elektronik- und Sensormaterialien, Technische Universität Bergakademie Freiberg

Martin Keyn, *Herstellung und Charakterisierung von Lanthanfluoridschichten*, Studienarbeit, 2010, Institut für Elektronik- und Sensormaterialien, Technische Universität Bergakademie Freiberg

Martin Keyn, *Schichterzeugung mittels Atomic Layer Deposition sowie spektroskopische und elektrische Untersuchung der Schicht*, Diplomarbeit, 2010, Institut für Experimentelle Physik, Technische Universität Bergakademie Freiberg

Betreute studentische Arbeiten

Bachelorarbeiten

Ba 267

Andreas Kramer,

AFM-Messungen an Kohlenstoff-Strukturen für die Nanoelektronik

Ba 268

Dennis Schäfer,

Tempereinfluss auf den Katalysator einer CNT-Herstellungstechnologie

Masterarbeiten

Ma 247

Johannes Palm,

Prozesstechnologie zur Herstellung von CNTFETs für Feldeffekt-Leistungsanwendungen

Ma 249

Dennis Noll,

Ellipsometrische und elektrische Charakterisierung von Katalysatorsystemen einer Herstellungstechnologie für Kohlenstoffnanoröhren

Danksagung

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am INSTITUT FÜR HALBLEITERTECHNIK UND NANOELEKTRONIK (IHTN) der TECHNISCHEN UNIVERSITÄT DARMSTADT.

Mein Dank gilt zuerst meinem Doktorvater, Herrn Prof. Dr. Udo Schwalke, für seine Unterstützung bei der Anfertigung dieser Arbeit. Herrn Prof. Dr.-Ing. Thomas Mikolajick danke ich für sein Interesse an meiner Arbeit und die Übernahme des Koreferats. Den Herren Prof. Dr.-Ing. Helmut Schlaak und Prof. Dr. techn. Heinz Koepl gilt mein Dank für ein anregendes Disputationsgespräch sowie Herrn Prof. Dr. Andy Schürr für die Übernahme des Vorsitzes der Prüfungskommission.

Herrn Dr.-Ing. Frank Wessely danke ich für seine Einführung in die Feinheiten der Halbleitertechnologie am Institut. Herrn Tillmann Krauss gilt mein Dank für *fruitful discussions* und die Unmengen an Kaffee, die währenddessen konsumiert wurden. Auch gilt mein Dank Frau Dr. Barbara Abendroth von INSTITUT FÜR EXPERIMENTELLE PHYSIK für die Möglichkeit, im Reinraum der TECHNISCHEN UNIVERSITÄT BERGAKADEMIE FREIBERG die Abscheidung des Aluminiumoxids mittels ALD durchzuführen zu können. Außerdem danke ich Herrn Andreas Kramer für die vielen elektrotechnischen Basteleien, die die elektrische Charakterisierung erleichtert haben. Schließlich danke ich allen Kollegen am Institut für die angenehme Arbeitsatmosphäre.

Lebenslauf

Name Martin Keyn
Geburtsdatum 15. Juli 1984
Geburtsort Grimma
Staatsangehörigkeit deutsch



Ausbildung

- **Schuljahr 1991/92 bis 1994/95:**
Theobald-Ziegler-Schule in Frankfurt am Main
- **Schuljahr 1995/96 bis 2003/04:**
Lessing-Gymnasium in Frankfurt am Main
- **Wintersemester 2005/06 bis Wintersemester 2010/11:**
Technische Universität Bergakademie Freiberg,
Fakultät für Werkstoffwissenschaft und Werkstofftechnologie
Studiengang: Elektronik- und Sensormaterialien
Abschluss: Diplom-Ingenieur

Berufliche Erfahrung

- **Oktober 2004 bis Juni 2005:**
Wehrdienst im Logistikbatallion 12 in Kühlshausen
Grundausbildung, danach Rechnungsführer-Gehilfe
- **Wintersemester 2008/09:**
Freiberger Compound Materials GmbH in Freiberg
Industriepraktikum im Rahmen des Studiums
- **März 2011 bis Februar 2017:**
Technische Universität Darmstadt,
Fachbereich für Elektro- und Informationstechnik,
Institut für Halbleitertechnik und Nanoelektronik
wissenschaftlicher Mitarbeiter und Promotionsstudent